OXFORD

Scientific Metaphysics

EDITED BY
Don Ross, James Ladyman,
and Harold Kincaid

Scientific Metaphysics

Scientific Metaphysics

EDITED BY

Don Ross, James Ladyman,
and Harold Kincaid

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© The several contributors 2013

The moral rights of the authors have been asserted

First Edition published in 2013

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Data available

ISBN 978–0–19–969649–9

Printed in Great Britain by
MPG Books Group, Bodmin and King's Lynn

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Acknowledgements

The editors extend their gratitude to Peter Momtchiloff of Oxford University Press for supporting and helping to refine the project; to Nelleke Bak for turning a ragtag mass of document files in different formats and reference conventions into a coherent manuscript; and to Susan Frampton, Sarah Parker, Prabhavathy Parthiban and Albert Stewart for technical production.

A previous version of chapter 7 was originally published in *Noûs* and is reprinted with permission.

Contents

Contributing Authors viii

1. Introduction: Pursuing a Naturalist Metaphysics 1
   Harold Kincaid
2. On the Prospects of Naturalized Metaphysics 27
   Anjan Chakravartty
3. Scientific Ontology and Speculative Ontology 51
   Paul Humphreys
4. Can Metaphysics Be Naturalized? And If So, How? 79
   Andrew Melnyk
5. Kinds of Things—Towards a Bestiary of the Manifest Image 96
   Daniel C. Dennett
6. The World in the Data 108
   James Ladyman and Don Ross
7. What Can Contemporary Philosophy Learn from our 'Scientific Philosophy' Heritage? 151
   Mark Wilson
8. Neo-Kantianism, Scientific Realism, and Modern Physics 182
   Michael Friedman
9. Some Remarks on 'Naturalism' as We Now Have It 198
   Mark Wilson
10. Causation, Free Will, and Naturalism 208
    Jenann Ismael

Index 237

Contributing Authors

ANJAN CHAKRAVARTTY is professor of philosophy at the University of Notre Dame, where he is a faculty member of the Graduate Programme in History and Philosophy of Science, and a former director of the Institute for the History and Philosophy of Science and Technology at the University of Toronto. He is the author of *A Metaphysics for Scientific Realism: Knowing the Unobservable* (2007), and editor of 'Explanation, Inference, Testimony, and Truth: Essays Dedicated to the Memory of Peter Lipton' (*Studies in History and Philosophy of Science*, 2010).

DANIEL C. DENNETT is university professor and Austin B. Fletcher Professor of Philosophy, and co-director of the Center for Cognitive Studies at Tufts University. He is the author of *Consciousness Explained* (1991), *Darwin's Dangerous Idea* (1995), *Freedom Evolves* (2003), and other books and articles, including 'Real Patterns' (J. Phil. 1991).

MICHAEL FRIEDMAN received his Ph.D. in philosophy from Princeton University in 1973. After teaching at Harvard University, the University of Pennsylvania, the University of Illinois at Chicago, and Indiana University, he is now Frederick P. Rehmus Family Professor of Humanities at Stanford University. His publications include *Foundations of Space-Time Theories: Relativistic Physics and the Philosophy of Science* (1983), *Kant and the Exact Sciences* (1992), *Reconsidering Logical Positivism* (1999), *A Parting of the Ways: Carnap, Cassirer, and Heidegger* (2000), and *Dynamics of Reason* (2001). He is the editor (and translator) of *Immanuel Kant's Metaphysical Foundations of Natural Science* (2004), the co-editor (with A. Nordmann) of *The Kantian Legacy in Nineteenth-Century Science* (2006), and the co-editor (with R. Creath) of *The Cambridge Companion to Carnap* (2007).

PAUL HUMPHREYS is professor of philosophy at the University of Virginia. His principal research interests are in the philosophy of science, epistemology, and metaphysics, with a current emphasis on computational science and emergence. He has also published on explanation, causation, probability, and the intellectual effects of contemporary technology. His most recent book was *Extending Ourselves* (2004).

JENANN ISMAEL is a philosopher of physics. She has written on probability, time, causation, modality, and quantum ontology. She has also engaged with issues in the philosophy of mind, applying the formal apparatus associated with symmetry to

understanding interactions among perspectives, self-location, and consciousness. Recently she has been thinking about the relationship between physics and free will. She has been a professor at the University of Arizona since 1996 and held fellowships at Stanford, the National Humanities Center, and a QEII fellowship from the Australian Research Council.

HAROLD KINCAID is professor, School of Economics, University of Cape Town. He has published multiple papers on reductionism, scientific realism, inference to the best explanation, and causation. Among his books are *Philosophical Foundations of the Social Sciences: Analyzing Controversies in Social Research* (1996) and *Individualism and the Unity of Science* (1997), and the co-edited volumes *The Oxford Handbook of the Philosophy of Economics* (2009), *What is Addiction?* (2010), *The Oxford Handbook of the Philosophy of Social Science* (2012) and the forthcoming *Mental Kinds and Natural Kinds*.

JAMES LADYMAN studied pure mathematics and philosophy at the University of York, and then took a masters in history and philosophy of science and mathematics at King's College London. He completed his Ph.D., on the semantic approach to scientific theories and structural realism, under the supervision of Steven French at the University of Leeds in 1997. He has been assistant, deputy and co-editor of the *British Journal for the Philosophy of Science* and honorary secretary of the British Society for the Philosophy of Science. He is professor of philosophy at the University of Bristol.

ANDREW MELNYK is a professor of philosophy at the University of Missouri. He has published papers and reviews in the *Journal of Philosophy, Noûs, Mind, Philosophy and Phenomenological Research, Philosophy of Science,* and the *Australasian Journal of Philosophy*. He is the author of *A Physicalist Manifesto: Thoroughly Modern Materialism* (2003), which develops and defends a comprehensively materialist view of the world.

DON ROSS is professor of economics and dean of commerce at the University of Cape Town, and section Director (Methodology), and research fellow in the Center for the Economic Analysis of Risk at Georgia State University. He holds a Ph.D. in philosophy. His current research areas are economic methodology, the experimental economics of impulsive consumption (e.g. addiction), strategic aspects of the evolution of human sociality, the promotion of scientism in philosophy and elsewhere, and relationships between risk and development in Africa. He is the author or editor of fourteen books and numerous journal articles and book chapters. Some recent working papers and pre-published versions of work are available at <http://uct.academia.edu/DonRoss>.

MARK WILSON is professor of philosophy at the University of Pittsburgh and a member of the American Academy of Arts and Sciences. His main research investigates the manner in which physical and mathematical concerns become entangled with issues characteristic of metaphysics and philosophy of language (his 2006 OUP book, *Wandering Significance,* explores such topics in detail). He formerly supervised the North American Traditions Series for Rounder Records.

1

Introduction: Pursuing a Naturalist Metaphysics[1]

Harold Kincaid

1. Introduction

This volume is about the prospects for a naturalized metaphysics and its relation to traditional metaphysics. One overarching theme is that traditional metaphysics, especially in its current incarnation as analytic metaphysics, is a questionable enterprise because of its lack of scientific standing. The thesis is that any legitimate metaphysics and conceptual analysis must be tied into the results and practices of the sciences. Thus a key motivation behind this volume is to explore what a metaphysics looks like that is judged by scientific standards and that avoids appeals to intuition. Many of the chapters are by authors who have written recent books on these issues and thus this volume is part of an ongoing debate that those volumes have started. The chapters are contributions that can be understood independently of that larger debate, but they are also an invitation to engage in it.

This chapter is divided into three parts. The first part sets the stage by outlining in general terms what form of 'scientific philosophy' or naturalism is at issue in most of the chapters. The second outlines the books and debates that this volume aims to extend. The third surveys the chapters and major themes across them.

[1] I would like to thank Don Ross for very helpful comments on previous drafts of this chapter.

2. Scientific philosophy and naturalism

Scientific metaphysics is in the broader tradition of 'scientific philosophy'. The term 'scientific philosophy' is borrowed from Reichenbach. For him scientific philosophy was using scientific methods and results to solve problems arising from science that are roughly philosophical in nature. This certainly describes the attitude and practice behind most of the chapters in this volume. Scientific philosophy is a version of naturalism. However, not everyone in this volume would call themselves naturalists, in large part because so many different and often unclear theses have fallen under that label. However, naturalism is central to this volume, and some delineation of specific relevant naturalist theses will help give more content to the idea of scientific philosophy and provide a deeper background for the chapters that follow.

Most versions of naturalism can be distinguished from one another by their answers to questions like the following:

- Is naturalism an epistemological claim about how we know or an ontological claim about what exists (or both or neither)?
- By scientific, are we talking about science as a method or science as a set of results or findings, or both?
- Is science only the non-human sciences or does it include the social and behavioural sciences as well?
- Is a priori knowledge about objective reality possible?
- What is the status of conceptual analysis and intuitions?
- What is the relation between philosophy (in our case, metaphysics) and science? More specifically, is philosophy constrained in any way by science and, if so, what does the constraint consist in?

The answers to some of these different questions are at least to some extent logically independent, so different subsets of them may get different answers. As a result, numerous different naturalisms emerge of quite varying logical strength, thus explaining why so many doubt that there is a meaningful doctrine to be affirmed or denied. However, the existence of many different possible formulations need not entail that no specific formulation is coherent and interesting.

There are self-avowed naturalists who answer these questions in ways that result in very thin naturalisms. For example, de Caro and Macarthur (2004b) defend a naturalism that holds that science can be relevant to

philosophy. They do not specify whether they mean that scientific results and/or scientific methods can be relevant to metaphysics. However, it is clear from other views they endorse that they reject the claim that scientific methods can solve metaphysical problems (not to mention the stronger claim that they are the only basis for doing so). Many contemporary metaphysicians on the other hand claim that they are naturalists because metaphysical results can have the virtues typical of scientific reasoning—simplicity, fruitfulness, and so on. (Criticisms of this move are raised below.) For instance, McDowell (2004) describes himself as a 'liberal' naturalist because thinking and knowing are 'concepts of occurrences and states in our lives', yet his putative 'naturalism' takes understanding human beings outside the standard domain of the natural in that human phenomena call for Verstehen, not reference to the laws and causes of natural sciences. Yet others count themselves as naturalists simply because they reject the supernatural.

The kind of scientific or naturalistic metaphysics that motivates this volume is considerably more demanding than these forms of naturalism, though no one exact formulation is needed or presupposed by all the chapters. Common to nearly all of them, however, are 1) an extreme scepticism about metaphysics when it is based on conceptual analysis tested against intuition, and about any alleged a priori truths that such intuitions and analyses might yield; and 2) the belief that scientific results and scientific methods can be successfully applied to some problems that could be called metaphysical. The conjunction of these two theses then provides some pressure for the stronger view that it is only by means of scientific results and scientific methods that metaphysical knowledge is possible, for it is not clear what third activity metaphysics might be if it is not conceptual analysis or scientifically inspired metaphysics.

To be explicit about what kind of metaphysics and views about philosophy's reach naturalism rejects, let me cite some explicit statements of approaches to metaphysics that are not naturalist:

1. The a priori status and priority of philosophy over science is defended by Bealer in a series of articles in prestigious analytic journals (Bealer, 1992, 1996a, 1996b, 1999, 2000). He argues for a) the 'Autonomy of Philosophy', which consists in the fact that most central questions in philosophy can 'in principle be answered by philosophical investigation' (1996a: 121) without relying on the sciences; and b) the

'Authority of Philosophy', which says that in cases of conflicts between science and philosophy the authority of philosophy is greater. The main article defending these views is cited 136 times in Google Scholar and a main application of his approach (Bealer, 1982) rates 416, quite good rates for work in philosophy. Of course, while all those citations could be from critical naturalists, a quick look at the individual citations does not support that interpretation over the conclusion that these ideas are taken very seriously by philosophers.

2. John Searle (2009: 9) expresses similar views about social ontology and its relation to social science. According to him, 'social ontology is prior to method and theory' in that philosophical questions about the nature of the social have to be answered first before any social science can be done and doing good social science requires having the correct philosophical analysis.
3. Lowe (2002: v) in his *Survey of Metaphysics* says that 'metaphysics goes deeper than any merely empirical science, even physics, because it provides the very framework with which such sciences are conceived'.

Aside from these explicit statements, many analytic metaphysicians in practice are non-naturalist in that their metaphysical claims are incompatible with the results of our best-established science. Ladyman and Ross (2007, Chapter 1) detail a number of such cases, and Humphreys in this volume argues that David Lewis's (1994) Humean supervenience is in flat contradiction to results in physics.

Put positively, the scientific philosophy and naturalism motivating this volume provides the following answers to the questions raised above. Scientific philosophy or scientific naturalism makes both epistemological claims and ontological/metaphysical ones. The epistemic claim is that all knowledge—including what we know about knowledge itself—comes from application of the broad methods and results of the sciences. Results and methods are intertwined because scientific methods may produce results that give us further information about methods. However, what counts as scientific methods and scientific results is not determined by a logic of science that demarcates science from non-science, for that would reintroduce an unwanted and unwarranted a priori foundationalism. As we will see in discussing Maddy's *Second Philosophy* in the next section and in the discussion of recurring criticisms of naturalism immediately below, the

kind of naturalism inspiring this volume rejects various anti-naturalist approaches that depend on empirically unwarranted assumptions about knowledge itself. And, just as there are no a priori constraints on what is science, scientific methods and results can characterize areas not normally thought of as scientific, for example, forensics, auto mechanics, and common-sense reasoning about causality. There also is no reason to restrict science to the natural sciences and exclude the social and behavioural sciences. Indeed the latter can be quite useful in investigating knowledge production in the natural sciences.

It follows that on the scientific naturalism espoused here standard a priori metaphysical knowledge is a nonstarter as is traditional conceptual analysis as a route to metaphysical knowledge all by itself. However, a naturalized form of a priori knowledge is not precluded, perhaps in the form of reliable intuition as defended by Anthony (2004) and discussed by Humphreys' chapter in this volume or in the form of a relativized a priori about the necessary preconditions of specific theories, as advocated by Friedman in Chapter 8.

As an ontological claim, scientific naturalism says that we know what exists through the application of scientific methods and results as understood above. However, scientific studies of science themselves may show that deciding what exists is not a matter of simply reading off the necessary prerequisites of scientific theory. In particular, Wilson's results discussed in the next section and in his two chapters in this volume make this prospect particularly unlikely.

To further flesh out the naturalist perspective, I want to turn next to answer some common charges against naturalism. The first is that contemporary analytic metaphysics does have naturalist virtues (Manley, 2009), or put in another way, that naturalism makes such minimal requirements that it threatens to be uninteresting (Chakravartty, this volume Chapter 2). There is a strong strain in analytic metaphysics that argues for or, more commonly uncritically presupposes, a broad notion of evidence, generally explicitly appealing to Quine's holism, and asserting that because this broad notion is the correct one, metaphysical claims can have the same scientific status as any other claim.

It is worth spelling out these ideas about evidence in more detail, both because of their role in supporting analytic metaphysics and because they are an important historical strand in the scientific philosophy that motivates this volume. The broad notion of evidence rehearses two familiar

Quinean[2] themes: that our beliefs form a web that faces the tribunal of experience as a whole and thus that any belief is ultimately revisable in the face of recalcitrant observations and that theoretical virtues, sometimes equated with pragmatic virtues, are part of our evidence and are evidence over and above the information impinging on our sensory surfaces, as Quine would put it. The first theme argues that our beliefs form a network such that there is always logical, evidential, or conceptual connections, albeit indirect, between any two beliefs, no matter how far apart they are in the web. This claim about evidence is often allied with or allegedly supported by parallel semantic theses sympathetic to the idea that meaning is use or is determined by the functional role in an overarching system. The second idea that theoretical or pragmatic values count as part of our reasons for belief is perhaps less explicitly developed, but it is motivated by (and perhaps an argument for) the key Quinean idea that theory is underdetermined by observation.

This broad conception of evidence has allowed many to pursue analytic metaphysics while claiming Quinean naturalist credentials. Different metaphysical views apparently can be rated in terms of their theoretical virtues, for example, in terms of their simplicity. Moreover, the argument goes, metaphysical claims are just at another level of abstraction beyond those of theoretical science, and like the latter, ultimately have ties in the web of belief to observation. Metaphysics is the attempt to see how knowledge in general hangs together and metaphysical theories can differ in their explanatory power, simplicity, and overall cohesion with the rest of our beliefs (including scientific ones).

Historically, these ideas have deep roots in mature positivism, Carnap in particular. As Friedman (1999) argues persuasively, Carnap did not advocate trying to ground science in the certainties of sense data and he was quite sympathetic to the role of pragmatic and theoretical values in science (Irzak and Grunberg, 1995). Science, according to Carnap, operates with 'frameworks' that are not built up by ostensive relations to the given but rather proceed top down from the structural relations between fundamental

[2] The use of 'Quinean' here means that I am not asserting that in the end this is the correct or most plausible interpretation of Quine's final views. He seemed to back away from extreme holism in asides written late in his career and the relation between theory underdetermination and the indeterminacy of translation in Quine has long been a matter of debate.

concepts. Frameworks are to be made explicit by philosophy of science using logic, set theory, and centrally notions of implicit definability, producing axiomatic formalized accounts of scientific theories. Philosophers can pursue this project on their own once they acquire the requisite scientific concepts. (This assumes what Wilson below calls the 'classic' conception of concepts.) So both the holism of evidence, its related semantic doctrines, and the place of pragmatic virtues in theory choice preceded Quine. Indeed, as is clear from Creath (1991), Quine elaborated these views in direct intellectual and personal engagement with Carnap.

This broad view of evidence is, I believe, misleading, and many chapters in this volume explicitly or implicitly reject it. It is important to distinguish two separate claims associated with the 'holism of testing': (1) the idea that every test of a hypothesis relies on background knowledge and theory and (2) the thesis that every test is and only is a joint test of the whole of theory, background knowledge, and so on. The second kind of claim is the one used to argue that metaphysics in traditional guises is tested against experience and has scientific virtues.

However, the stronger holism does not follow from (1) and is in fact quite implausible. As Glymour (1980) pointed out some time ago, it is possible to use bootstrapping and triangulation procedures to localize blame for empirical failures. Science makes such arguments all the time. Furthermore, there is reason to be suspicious of theoretical or pragmatic virtues such as simplicity and inference to the best explanation. Sober (1988) argues that appeals to simplicity in phylogenetics are really disguised empirical hypotheses about the nature of evolutionary processes. Day and Kincaid (1994) argue that fruitful inferences to the best explanation are really implicit appeals to substantive empirical assumptions, not to some privileged general form of inference. Ladyman and Ross in Chapter 6 point out that inference to the best explanation is often just inference to the most intuitive explanation, making metaphysics again an unwarranted conservative constraint on science. Likewise, Friedman in this volume provides another reason to doubt the extreme holism embodied in thesis (2): fundamental physical theories can be divided into empirical claims and the coordinating framework that makes them possible. The epistemic status of the two is quite different.

The kind of metaphysics defended in this volume is one that meets considerably more stringent conditions than the 'hazy holism' (Wilson's phrase) of Quine as manifested in thesis (2). Humphreys argues that science

certainly asks ontological questions ('Are there Higgs bosons?'[3]), but that the evidence and criteria for existence claims is often domain specific and thus that general criteria such as Quine's distort the science. Wilson argues that the actual content of scientific concepts is much too open-ended and application specific to allow talk of 'the' formalizable theory in many areas of science. (He gives many examples from the mechanics of medium-sized, medium-distanced objects.) So testing whole theories by the total evidence is a mistaken project and distortion of what science does. If there is metaphysics in science, it will have to be quite local. Ladyman and Ross in particular argue that legitimate metaphysics must meet a quite tight requirement—it must show how to unify apparently independent scientific claims.

Another common objection to naturalism (Williams, 2001) is that it leaves no place for normative epistemic judgements since it replaces epistemology with purely descriptive science. The fallacy here is to think of science as purely descriptive. Scientists are essentially worried about what we are entitled to or ought to believe, and much of what they do is aimed at establishing evidence and reasoning about those obligations and entitlements. The reliability of methods, new and old, is a key scientific question towards which great attention and resources are directed. Even statistics, which one might think is all about implications from mathematical axioms about probabilities and distributions, is subject to this kind of scrutiny—the reliability of new statistical procedures these days is often determined by checking them against known, simulated data to determine their sensitivity and specificity, not by proofs of asymptotic characteristics. These naturalized approaches are not restricted to citing established facts about psychological processes in the way Quine seemed to suggest, and even those facts can have normative implications—witness the heuristics and biases research programme in cognitive science and economics.

A further standard charge against naturalist approaches is that they are question-begging—what grounds naturalist epistemological claims themselves? The threat of circularity allegedly shows that that there must be non-naturalist aspects of evidence and justification. A related worry is that naturalist epistemology is simply a subject-changing exercise to the extent that it dismisses issues of scepticism. There is a long diagnostic story to be

[3] Ladyman and Ross use as their example 'Is there sex addiction?'

told here about why these claims are misguided, and the story's best teller in my view is Williams (2001). A shorter story will have to do here. A sophisticated naturalist denies that there is an informative, nontrivial set of scientific methods that the naturalist then applies to philosophical problems and which themselves need justification—this presupposes a logic of science ideal that disciplined inquiry into the ways of scientific practice show is untenable. Conceptual and metaphysical problems arise in the practice of science and can be addressed in the diversity of ways through which science proceeds, above all by the filtering processes of institutionalized peer review. A sophisticated naturalist also denies that there is a particular philosophical method and standpoint from which to make a priori judgements about the requirements for knowledge. Sceptical arguments get their purchase by making assumptions about knowledge that do not fit with our scientific and everyday knowledge practices. Those practices always involve challenges against a background that gives them sense and makes them possible. 'Our knowledge of the external world' and 'adequately grounded beliefs' are philosophical simplifications that are unmotivated by real science and real knowledge production. Quine's naturalism unfortunately smuggles these simplifications back in by invoking 'stimulations of sensory surfaces' and a web of belief held together by deductive relations. While local sceptical arguments can have an important place in science, as Ladyman and Ross urge in Chapter 6, global arguments presuppose situations we are not in and conceptions of knowledge that are unmotivated.

3. Summary of previous volumes

As noted earlier, many of the chapters in this book extend arguments that appeared in or were inspired by books published within a short of time of each other and that approach issues in science and metaphysics in related but different ways. Friedman (2001, but also see 1992 and 1999), Ladyman and Ross (2007), and Wilson (2006) are precursors to the ideas these authors defend in their chapters and motivate the topics of others. Maddy (2007) is another leading defence of the kind of naturalism sympathetic to the projects pursued by our contributors. This section outlines the main ideas of these volumes in so far as they are relevant.

Friedman's work actually rests on serious and ground-breaking historical scholarship on Kant's project. In *Kant and the Exact Sciences* (1992) Friedman argues that Kant's work cannot be properly understood apart

from the problems in the natural sciences he was trying to clarify. Working from Kantian texts in the pre-critical period, from Kant's *Metaphysical Foundations of Natural Science* from the critical period, and from discussions of chemistry in the postcriticial posthumous writings, Friedman shows that Kant is fundamentally concerned to find an account of metaphysical ideas[4] suitable to science and to explain their role in science, mainly in the specific context of the Newtonian system that Kant clearly expected to endure as the permanent basis of fundamental physics. That role is one of supplying necessary presuppositions or the definitive constitutional principles that make the empirical application of Newtonian principles possible. Combine this neglected side of Kant with the well-known Kant who thinks that traditional metaphysics—metaphysics that is not tied to the possibilities of experience—leads to absurdities and we almost have Kant the naturalist doing 'scientific philosophy', motivated by and helping to solve problems in the sciences.

Friedman's *Dynamics of Reason* (2001) builds on this interpretation of Kant. Friedman brings to this book his understanding of developments in relativistic physics since Kant and a convincing picture of what the positivists were really trying to do (in his 1999 book). On Friedman's view positivists such as Carnap were in effect neo-Kantian in that they left some place for the a priori as the necessary framework components of physical theory and saw philosophy's job through formalization using logic and set theory as identifying those elements. Using a distinction found in Reichenbach and Carnap, he distinguishes two kinds of possible a priori truths—those that are known independently of experience, unrevisable, and certain and those that are essentially presupposed by empirical work by constituting and defining the objects of investigation. Friedman defends only the latter sense of a priori principles and works out the differences in their content from Newton to Einstein. Newton's laws of motion made universal gravitation an empirical claim by implicitly defining the relevant spatiotemporal structure. Something similar can be said about Einstein and general relativity: non-Euclidean geometry and tensor mathematics combined with Einstein's principle of equivalence are the necessary preconditions making Einstein's field equations possible. These assumptions

[4] Kant himself reserved the term 'metaphysics' (in its German equivalent) for the traditional varieties he rejected. Friedman follows this preference with respect to contemporary usage.

about spatiotemporal structure are metaphysical ones and ones partly drawn from traditional metaphysical debates. The resulting constitutive principles are revisable and subject to empirical refutation in that empirical results may lead to changes in the framework of constitutive principles as theories change. However, they are not subject to direct empirical investigation and Friedman breaks with the Quinean doctrine that all parts of a theory have the same epistemic status.

For Friedman this relativized a priori is what provides objectivity in science understood as intersubjective agreement. The metaphysical constitutive principles identified tell us about the nature of our *representations* of the world, not the world in itself as it were. Different frameworks for Friedman are in a sense similar to Kuhnian paradigms, but Friedman argues that the interaction of scientific philosophy as done by Mach, Helmholtz, and Poincare, along with developments in the sciences helped make the moves from Newton to Einstein possible and understandable (thus rejecting Kuhn's strong incommensurability). Friedman's interpretation of metaphysics as being about our conceptual structures, not independently existing objective reality, makes his stance on metaphysics quite different from those of Ladyman and Ross, Ismael, and Humphreys who defend the possibility of metaphysics as knowledge of that objective world.

Friedman's scientific philosophy shares a theme common to most chapters: philosophy can be a source of creativity and added value to science if it arises from engagement with real scientific problems. When it does not and imposes metaphysical concepts from refined common sense, it can be a positive hindrance.

Mark Wilson's *Wandering Significance* (2006) is a study of how concepts work in fine detail, with concepts from macrolevel physics and applied mathematics being the prime examples. Wilson criticizes a classical (a la Russell's *Problems of Philosophy*, 1912) account of concepts that he believes still holds considerable sway in philosophy. The classical view holds that once one has acquired a concept, its contents are relatively transparent. The view also tends to depict concepts as fixed entities. Wilson shows the motivations of such a conception in developments in nineteenth-century science and then through detailed case studies shows how scientific practice has undermined the classical picture of the concept and in fact cannot be understood without rejecting it. In particular he shows that classical mechanics is a 'patchwork' of theory 'facades'. These are successful inferential practices that vary from application to application and that are built around

mini-theories that need not be mutually consistent when extended outside of their domains of application, making axiomatizations or Kuhnian paradigms of doubtful value for understanding science. Moving from patchwork to patchwork without error requires careful instructions and procedures learned in practical apprenticeship. Wilson illustrates such procedures in great detail, turning the vague idea of meaning as use into a study of science in practice that details the various strategies that make concepts work and allows them to adapt to new circumstances. The account he gives shows that there are serious real world constraints on what can be done, something he thinks a Quineian 'meaning as position in the web of belief' misses. Thus he shares with Friedman a denial that all scientific propositions have the same status in terms of connection to observational evidence and that they do so only along with all the rest of the propositions with which they are inferentially and evidentially related.

Wilson's book is not directly about science and metaphysics, but it does have significant implications for metaphysics once a naturalistic approach is presumed. Analytic metaphysics relies very much on a classical conception of concepts. Wilson shows that important concepts cannot be adequately understood just by intuiting the sense that comes from having the concept. Standard conceptual analysis is ill-equipped to handle the kind of complexity Wilson identifies. Wilson does conceptual analysis to be sure, but in a way that makes for a much more rigorous and enlightening project than testing necessary and sufficient conditions against what philosophers would say or find intuitive. In another sense, Wilson is doing metaphysics: he is showing how fundamental features of such concepts as force or even concepts for secondary properties hang together.

However, Wilson's work is anti-metaphysical in other ways. The concepts he studies are a rigorously formulated and extended in the context of empirical inquiry. This is usually scientific inquiry, but given Wilson's focus on applied science, it need not be fundamental physics. For example, to get a grip on 'red' it is helpful to look at what photographers and professional printers know about the use of colour terms (which is quite complicated). Wilson doubts that many of the common-sense concepts that are used in standard metaphysics—of the 'doorknob' and 'dog' variety—have the kind of rigorous tie to empirical data and mathematical formulation that allows them to help us to make real progress in understanding how the world works. It is useful here to invoke a distinction common now in ethics between thick and thin concepts, with 'cowardly' representing the former

and 'good' being an instance of the latter. Wilson can be seen as arguing, as do others who invoke this distinction, that the real work is done by the thick concepts, that that is where our philosophical attention should be focused, and that it is in applied scientific usages that many of our initially thin concepts get their thick content. ('Hardness' is one of his leading examples.) Wilson seriously undermines the account of concepts typical in traditional metaphysics and recommends a replacement activity that address real scientific problems with some kin to broad metaphysical concerns.

The recent books by Maddy (2007) and Ladyman and Ross (2007) extend this debate by defending a more explicit place for metaphysics within a naturalistic philosophy of science. Maddy provides a sophisticated defence of scientific philosophy with applications to metaphysics; Ladyman and Ross produce a critique of traditional analytic metaphysics from a naturalist perspective as well as a defence and instantiation of metaphysics properly motivated and constrained by scientific problems.

Like Friedman and Wilson, Ladyman and Ross want conclusions about the status of metaphysics to be motivated by scientific results, scientific problems, and scientifically inspired methods. They think contemporary analytic philosophy rests on multiple mistakes. They point to core naturalist complaints about conceptual analysis—that it is a priori, that it relies on the intuitions of philosophers that are suspect, that our best empirical evidence is that important concepts are not captured by necessary and sufficient conditions, that common sense ideas can and should be overturned by established scientific results, and so on.

Beyond these complaints, they claim to show that contemporary analytic metaphysics is often in direct contradiction with well-established scientific results. Their most dramatic claim is that modern quantum physics shows that there are no things, only structures—they defend ontic and epistemic structural realism—and so metaphysical questions about the identity and persistence of statues or persons, for example, are beside the point. Fundamental physics is not about little things banging into each other. Relatedly, fundamental physics does not support the containment metaphor behind so many debates about composition and mereology in contemporary metaphysics.

Ladyman and Ross combine these critiques with a positive explanation for many of our ordinary practices, especially much of the special sciences, that seem to be about things. Things and individuals are useful

bookkeeping devices for identifying what they call—following and refining Dennett (1991)—real patterns. Real patterns are based on the ability of sources of Shannon information to resist entropy. For humans with limited cognitive resources various thing concepts are quite useful for picking out those real patterns using parameters that are not from fundamental physics and identify only a restricted aspect of the universal structure identified by the latter.

Ladyman and Ross go on to develop in detail a defence and picture of scientific metaphysics to replace the view they reject. There is a legitimate place for metaphysics in unifying the diverse results of the special sciences with fundamental physics. This is an activity that may generally fall between the cracks of the sciences, so it is a useful enterprise. It is also an activity that must be constrained by the institutionally produced results and methods of the sciences. An additional constraint on such unification added by Ladyman and Ross is that the special sciences cannot contradict or override the results of fundamental physics.

Maddy provides one of the best recent illustrations of naturalism. (She does not call her book a defence, mistakenly in my view—see below.) She describes a character she calls the Second Philosopher, the way this philosopher proceeds in general, and how she might approach issues in philosophy of mathematics and metaphysics. Second Philosophy is of course contrasted with First Philosophy. First Philosophy is in Maddy's account of the attempt to give extra-scientific foundations to the sciences and to other kinds of knowledge as well. While she declines to define second philosophy, the denial of first philosophy goes some way in doing so—second philosophy is the attempt to solve problems, including what get called 'philosophical' ones, using the broad methods of the sciences. I say 'broad' because Maddy wisely declines the project of giving necessary and sufficient conditions for 'is science' or 'is a scientific method'. The Second Philosopher does not need to appeal to the authority of something called science but instead appeals to available scientific evidence and methods. Maddy explains why a Second Philosopher is not moved by skeptical arguments and views like those of Kant, Carnap, Putnam (in one guise) and others who claim to have a two-level or two-perspective view of science—one from within science and one from without. She denies that the skeptical position is just a generalization of the standard problems of knowing in common life and its refinements in science. She also does not see the pull of giving a perfectly general account of how our

knowledge or language relates to the world and nor does she believe that, pace Quine, determining what there is can be simply read off our successful scientific theories. In these regards she cites Wilson (2006) as a fellow traveller and argues that what science is committed to at any given time is a complex matter and that it is a mistake to regard the ontology of science as something postulated as the 'best explanation' of sciences themselves, for the latter presupposes an external viewpoint she sees no reason to take. She also does not take 'science says so' to be a good form of argument, but instead prefers the scientific evidence itself. Metaphysics thus on her view consists in case studies of scientific results that help us case by case 'to reveal and explicate the rationality of the existence claims based on each of these [scientific] particular varieties of evidence' (Kindle location 4468), something we can take Wilson to be doing in part. This project can be called philosophy done within science, but the important point is that it is within science.

Maddy denies that she is defending second philosophy against traditional philosophy, saying that the concerns of the two are different. But she is too modest, on my view. She is giving reasons, based on common everyday knowledge acquisition processes and their refinement in science, for thinking that the demands of traditional philosophy are unreasonable, especially analytic epistemology. As I suggested earlier, Williams has made this kind of case with considerable power and detail.

4. The chapters

Chakravartty's chapter is helpful in asking what naturalistic metaphysics might be and what might motivate it. He is sympathetic to what he takes to be the main motivation, viz. the great distance between analytic metaphysics and the empirical results of science. However, he doubts that there is currently a coherent and clear sense of naturalistic metaphysics, arguing that we cannot keep a priori metaphysics out of science because metaphysical notions such as cause inevitably arise in reporting scientific results; because there is at present no clear, non-trivializing way to distinguish metaphysics that is based on the a posteriori part of science from that which is not; and because metaphysical questions are underdetermined by the evidential base of science. Chakravartty ends with a positive first stab at characterizing the sense in which metaphysics might be 'grounded' in empirical science, a standard but unclear naturalist claim. He spells out the idea in terms of

distance from experience running along a continuum from directly detectable on one end to inferred or postulated on the other and in terms of Popperian riskiness. Many of the remaining chapters can be seen as providing responses to Chakravartty's position.

Humphreys focuses on the status of what he calls 'speculative ontology' versus 'scientific ontology'. Scientific ontology gives us answers to questions about what exists that are consistent with the well-established results and methods of science. Speculative ontology is not so constrained. His main target and the bulk of his chapter is the latter.

Humphreys identifies multiple problems with speculative ontology. First he claims that it is sometimes directly inconsistent with well-established science. Since the science also is readily accessible and could be understood by philosophers, this kind of metaphysics is epistemically 'irresponsible'. A second problem is the reliance on intuition in speculative metaphysics, since intuitions differ and have limited reliability. Where we have conflicting intuitions and no criteria to determine who are the experts, intuitions should be rejected as evidence in general, and thus as evidence in metaphysics. Echoing a common theme of the volume, Humphreys also argues that any analysis of common-sense concepts cannot legitimately place constraints on science but speculative ontology threatens precisely that. Finally, Humphreys suggests that we have reason to believe that the world is not scale invariant (a theme emphasized by Ladyman and Ross as well), making our ordinary concepts and intuitions of dubious projectability beyond the domain of our own commonplace interactions with and manipulations of the world.

Melnyk in his chapter carefully sorts out key issues around what a naturalistic metaphysics would be like and what motivates it. A naturalistic metaphysics might show how claims of the special sciences hang together with fundamental physics as claimed by Ladyman and Ross, but why must it be restricted to unifying (as opposed to answering other questions), to identifying when unification is properly at issue, and to establishing unity with physics rather than with other parts of science?

Melnyk, like Humphreys, thinks naturalized a priori knowledge has not been ruled out. Intuitions may be reliable in some domains of inquiry. He also asks 'What sort of priority does science have?' The common-sense practices for finding out about the world that science refines may nonetheless still tell us some things about how the world is and so produce knowledge, at least where it is does not contradict well established

scientific results. Melnyk also asks whether a properly naturalized metaphysics could correct science, since on its advocates' view it is part of science itself. These are good questions.

Dennett proposes that we think of standard analytic metaphysics as sorting out the common view of the world which he equates with Sellars's notion of the 'manifest image'—the set of beliefs that ordinary humans have formed in their evolution so that they can get about in the world. On his view what drives much philosophy is negotiating between the scientific image and the manifest image, and it would be helpful to have an anthropology of the latter. This requires bracketing truth in favour of a careful delineation of its contours, producing a metaphysics of the manifest image.

This position bears some resemblance to 'the Canberra project' in metaphysics, which sees conceptual analysis as about determining the outline of folk theories. The Canberra project then wants to ask what in the world, if anything, corresponds to the elements of those theories. Dennett, however, instead talks of the manifest image as a framework and invokes the idea of ontological relativity, trying to avoid the conclusion that the manifest image is false or in need of corrections. As he grants, this line of defense calls for further elaboration.

Ladyman and Ross focus their chapter around consideration of a recent book by Deutsch (2011). Deutsch, a physicist, develops a philosophy of science and metaphysics based on the goal of seeing how sciences as diverse as physics, the theory of computation, and evolutionary biology hang together. Thus he shares Ladyman and Ross's (2007) thought that metaphysics in the form of unifying different scientific hypotheses is a worthy scientific project. He argues against reductionism because scientific results show that macro-level phenomena are often best explained in their own terms and that the scientific criterion for existence is information-theoretic—to be is to be a quasi-autonomous flow of information. These ideas are quite close to those of Ladyman and Ross (2007) and provide a context for them to further explain and clarify those claims in their chapter.

However, Deutsch in the end argues for an interpretation of quantum mechanics that presumes that irreducible stochasticity is scientifically unacceptable. For Ladyman and Ross, this move brings dogmatic, non-naturalized metaphysics into science. They propose instead a Pierceian understanding of the world such that it is the inherent stochasticity of

reality that explains the success of our statistical methods. The world is the totality of non-redundant statistics. They apply these ideas to debates about the interpretation of quantum mechanics in a way that is quite original and sophisticated and I will return to it in concluding the chapter.

Wilson presents his critique of the classical view of concepts and draws out the implications for what philosophy can and cannot do. The classical view of concepts—as found in Russell for example—asserts that once one has acquired language and the concepts it embodies then one can determine conceptual contents by introspection. On Wilson's view much contemporary philosophy rests on an inchoate amalgam of classical and anticlassical themes. Scientific philosophy of the past was motivated by scientific problems that required a conceptual freedom and creativity that did not fit well with the classical conception of concepts. However the positivist tradition bet on a wrong alternative to the classical notion, depending on the idea that set theory, logic, and axiomization would supply a curative. A real understanding of ongoing scientific conceptual change in the last two centuries in mathematical physics, especially the physics relating mechanics to constituents of matter, shows that 'science comes at us in pieces' that invoke a variety of stratagems embodying applied mathematics, making for piecemeal local change to develop concepts that reliably track the data. As an illustration, Wilson discusses in detail the problem of explaining the motion of a bead on a string, Euler's rule for doing so, and the obstacles and changes needed to make the rule reliable.

Much conceptual and inferential change in science is of this sort, making Wilson suspicious of axiomatizations and talk of overarching frameworks in science. Thus he is doubtful about projects, like that of Friedman (2001), according to which modern day scientific philosophy should be involved in creating the conceptual space for understanding and transforming contemporary frameworks in science. He instead thinks the study of scientific concepts in practice in all their localized detail is the kind of work that someone trained in the analytic methods of philosophy can better undertake. This is a project far removed from the analysis of necessary and sufficient conditions based on armchair intuitions about meanings.

Friedman lays out his program for scientific philosophy as the identification and development of the constitutive representations of mathematical physics. Like Wilson, he thinks that Quine's holism encouraged overlooking important differences in the kinds of assumptions at work in scientific theories. He explains how identifying a priori elements of

scientific frameworks is necessary to fully understand them, helps make them possible, and provides for a kind of objectivity as intersubjective agreement via a relativized a priori.

Friedman argues that this project is orthogonal to the project pursued by Wilson. The kind of tension and change produced in trying to connect mechanics and matter theory were not ones that drove the historical change to relativity theory that is Friedman's main case study. Einstein was careful to bracket issues in matter theory of the sort described by Wilson as he developed the theories of special and general relativity. So there is room for the study of concepts in the context of the fundamental presuppositions of theoretical frameworks and for the patchwork-like atlases of the application of physical concepts that Wilson produces. It is an interesting and open question how kinds of scientific activity aimed at modeling different aspects of systems might be interrelated.

Two questions about Friedman's approach arise. Wilson in this volume doubts that there is such a unified framework in the first place. As he would put it, science comes at us in pieces as a patchwork of adjustments whose structure is not easily characterized as an n-tuple. Wilson's *Wandering Significance* is a sustained argument backed up by countless detailed examples for this claim. So it unclear whether Wilson can ultimately accept Friedman's proposed reconciliation of their two projects as orthogonal. A second question concerns what Friedman's view implies for the scope of metaphysics. The presuppositions of physics address issues about the nature of space, for example, but other metaphysical issues don't seem to have any direct connection.

Wilson in his second chapter turns to a variant of naturalized metaphysics that he finds implausible, but I would also argue that his criticisms tell against strands of analytic metaphysics as well. He objects to the idea that we can adequately describe the physical world with only the kind terms of basic physics and first-order logical operations on those predicates, supporting the prospect that mathematics might be eliminated. The problem is that much mathematics is tangled up in specifying those predicates themselves, and applying them to reality involves determining a variety of other constructed physical quantities on a domain by domain basis. So clearly Wilson is a naturalist who is not committed to common forms of physicalism. I think that his point generalizes to problems for analytic metaphysics, because innumerable debates about whether physical properties are fundamental in various (obscure) senses take it for granted that we know what we

mean by 'physical property' in the first place (and by 'biological,' 'mental,' or 'social' properties for that matter). Debates framed in these terms are off to a bad start from the very beginning.

Following up on a comment from Dennett in *Freedom Evolves* that his own account might be improved by incorporating recent work on causality from philosophy of science and computer science, Ismael discusses the scientific philosophy project in relation to the metaphysics of free will. On her view naturalistic metaphysics means there is no science-independent standpoint for inquiry and involves applying science to philosophical problems. For example, the naturalistic metaphysician takes her view of cause from science. The common-sense notion of cause may be of a force imposed from outside. Philosophers may think of the deterministic laws governing entire physical systems as determining their subsystems. But these ideas are undermined by our best scientific account of causal knowledge. Ismael takes our best understanding to be based on the exciting developments originating in Pearl (2010) and Spirtes, Glymour, and Scheines (2001). They show that by working back from an explicit causal model meeting certain assumptions it is possible to deduce the conditional and unconditional dependencies that should be seen in the data if the model is the correct one, and give a semantics for graphical causal models that illuminates seemingly different empirical approaches to causation in the social, behavioural, and biomedical sciences. Ismael believes that this work shows among other things that what causes we find depends on what we hold fixed, that we cannot get theories of local subsystems from dynamic laws that pertain to the whole, and that local modalities are richer than global ones. These understandings defuse the antecedent objection to basically compatibilist notions of free will.

Ismael points out that the work on causation that she invokes is not about conceptual analysis of our ordinary concepts but an explication of causal inference in science. Ismael also argues that the Pearl and Scheines et al. programme in graphical causal modelling sheds scientific light on the nature of causation. I might add that it does so by in part showing that what seemed to be diverse understandings of causation in scientific practice are in effect reflections of one common underlying framework. This is a nice example of Ladyman and Ross's interest in metaphysics as producing scientific unification, although it does not involve physics in they way they would want (and with which Melnyk disagrees in his chapter).

It is worth expanding on this reading a bit to help clarify the versions of naturalistic metaphysics that predominate in this volume. Hall (2007), in a recent article on structural equation models and associated graphical explications of causality, complains that they do not (could never?) handle causal preemption. However, this misunderstands the project encapsulated in graphical accounts of causal models. It is not to give necessary and sufficient conditions for our ordinary notion of cause. It is to give a unified account of how evidence is amassed for certain specified notions of cause that scientists antecedently understand. The causal evidence amassed is supposed to be conditional on an already specified causal model. The results generally are explicit that preemption is not handled by these methods—the standard faithfulness condition rules them out in that when preemption is present, incorrect inferences are made. The notion of causation that is best understood by these methods is that of individually sufficient independent causes that contribute to an effect (Kincaid, 2012b). Pearl and company do not claim to handle all possible types of causation—necessary causes or effect modifiers do not easily fit into the framework. It is an empirical matter as to which kinds of scientific causal claims their models capture. Conceptual analysis in the sense of getting clear on how apparently different scientific methods and results hang together does go on but it is a quite different enterprise from conceptual analysis based on intuitions about our common-sense notion of cause. The unification achieved is not necessary for those practices to proceed but it is a unification that increases our understanding of how different methods and the results they produce fit together. It is quite different from the analytic metaphysical approach to causation. It gives no reason to think that kind of metaphysics is essential to science.

I want to finish my discussion of the chapters by returning to Chakravartty's chapter and making more explicit how and to what extent the chapters that follow answer his doubts about scientific metaphysics. There are various different and complementary responses to the claim that traditional metaphysics is inevitable in science. Humphreys argues that traditional metaphysics in the guise of speculative ontology is unconstrained by scientific evidence and often flatly at odds with it, and is thus illegitimate. However, scientific ontology goes beyond just what scientists explicitly are committed to because sometimes interpretations are underdetermined by the data as, for example, in quantum mechanics. However, those interpretations are still constrained by well-established results and widely accepted

standards of scientific evidence, unlike speculative ontology. So traditional metaphysics in the sense of speculative ontology has no place in an objective scientific understanding of the world.

Thus the underlying point is that 'metaphysics' comes to several different things and understanding naturalized metaphysics requires staying clear about these differences. Friedman echoes this idea from another vantage point. The claims that derive from the necessary presuppositions required to tie theory to observation are quite specific and must be distinguished from other traditional metaphysical claims that do not play that role.

Ladyman and Ross have interesting things to say about the scientific ontology that Humphreys identifies with competing interpretations of the same data and that constitute one of Chakravartty's reasons for thinking metaphysics cannot be naturalized. In *Every Thing Must Go* Ladyman and Ross deny that the status of questions that go beyond any possible evidence is an issue for science and thus a naturalistically respectable issue. The argument is not that such questions are meaningless, as positivists sometimes maintained. It is that science itself prohibits pursuing such questions, for they are not possible objects of 'objective inquiry'. Some of the kinds of underdetermination Humphreys has in mind may fall into this category.

In their chapter in this volume Ladyman and Ross address the interpretation of quantum mechanics directly and add interesting new content to their view on these issues, I believe. They point out that scientists can and should be strategic antirealists of the instrumentalist stripe on some occasions by refusing to take a stand on interpretations. They point out that working with a formalism that is given no preferred physical interpretation can lead to useful, novel, and compelling evidence supporting the structure behind the formalism. Such is the case with quantum mechanics.

Ismael's comments about the progress in understanding causation coming from graphical approaches replies to Chakravartty's claim that because scientists make causal claims, they are inevitably involved in metaphysical theorizing. As I pointed out earlier, the work of Pearl et al. is not about finding a general conception of causation in non-causal terms tested against intuition and common-sense counterexamples. Their work is showing how initially apparently different scientific ways of identifying causality can be brought under one formalism and improved thereby. When scientists use the graphical approach to make causal claims, they are not taking a stand on the analytic metaphysics of causation, which is a debate that has little direct relevance to their empirical concerns.

5. Final thoughts on open questions

The chapters in this volume suggest numerous open questions worth pursuing. I conclude with a discussion of some of them that I find particularly interesting.

Chakravartty's question concerning how science constrains metaphysics is certainly an open one that can fruitfully be explored. However, I doubt that searching for a perfectly general account of 'is constrained by' is likely to yield much insight. Assuming that there is *a* general account to be given buys into the kinds of assumptions typical of analytic metaphysics that this book challenges. Ladyman and Ross provide one general way that metaphysics might be constrained, namely, by unifying the special sciences with fundamental physics. They think of unification in terms of Kitcher's argument patterns, but an interesting question is whether other notions of unification might be applicable. Furthermore, as Melnyk suggests, there presumably are other kinds of constraints that can count as evidence for metaphysical claims. Wilson's work showing that scientific concepts get tied down to reality in diverse, complex, and local ways should be true of the ways in which science can constrain metaphysical claims. Combine that implication with the idea advanced in places in this volume that reading the metaphysical commitments off scientific results is also a local and complex task and the moral would seem to be that many different detailed case studies of metaphysics in the practice of science are called for. The Ladyman and Ross discussion of quantum mechanics in this volume shows in part how such case studies can be done.

Other fruitful case studies might take up the issue of how the two different approaches to scientific theory advocated by Friedman and Wilson are related. Friedman thinks it important to identify the necessary presuppositions that make theories applicable to the world; Wilson emphasizes the diversity of ways a single body of theoretical work gets connected to reality. Friedman thinks these approaches are orthogonal, at least for the physical sciences they both focus on. How the two different approaches actually could interrelate in telling the history of physics is in interesting open question.

While this volume has been suspicious of conceptual analysis and appeals to intuition, as Melnyk, Humphreys, and Dennett point out in different ways, there may still be useful work to be done if the practice is refigured. One reframing is to think of analytic metaphysics as sketching the commitments of our common-sense concepts—autoanthropology as Dennett labels it. If I read Dennett correctly, I have to wonder why the standard practices of

definition and counterexample based on intuitions practised by analytic philosophers are the way to get at the manifest image rather than by using standard scientific procedures such as those that cognitive science—particularly cognitive anthropology—uses. Of course, Dennett makes extensive use of such tools in his *Freedom Evolves* (2003), which is a paradigm of naturalist metaphysics applied to the free will problem. One also wonders whether, if freed from the classical picture of concepts, conceptual analysis of Wilson's style might have more useful things to say about common-sense concepts or whether they are too untethered to disciplined empirical constraints for such accounts to be possible. Similar questions arise for concepts from sciences, such as sociology or linguistics, that lack any special 'anchoring' mathematical tools that are distinctively characteristic of their practice.

There is also surely much room for investigations into the ways in which common-sense metaphysical concepts have influenced different sciences and whether for better or worse. To take an example used in this book, it is arguable that a particular common-sense conception of causation is behind the graphical approaches to causation, namely, that of the mechanic intervening by replacing parts while leaving all else the same. It is an open question whether all the kinds of causation recognized by common sense or by the sciences can be understood in such causal terms (e.g. Cartwright, 2007; Kincaid, 2012b).

A final useful and deep question raised by this volume is about conceptual analysis and intuitions and their actual and proper role *within* science. It is fine to say that many concepts are not captured by necessary and sufficient conditions, but surely it can be good to have concepts that are so defined—when is that and how does the science work without them? Wilson's work certainly bears on this topic but case studies more directly pursuing these questions might be revealing. Defining what they call ontologies is a major preoccupation of biomedical and bioinformatic sciences. This literature has not been discussed in this volume, but it certainly uses set theory and other logical tools to get at the basic kinds of things in the domain in a way that would not be entirely alien to analytic philosophers. How that work relates to themes and theses of this volume is one of many open questions to be pursued.

References

Anthony, L. (2004). A naturalized approach to the a priori. *Philosophical Issues*, 14: 1–17.

Bealer, G. (1982). *Quality and Concept*. Oxford: Oxford University Press.
—— (1992). The incoherence of empircism. *Aristotelian Society*, supplementary volume 66: 99–138.
—— (1996a). A priori knowledge and the scope of philosophy. *Philosophical Studies*, 81:121–42.
—— (1996b). On the possibility of philosophical knowledge. *Philosophical Perspectives*, 10: 1–34.
—— (1999). A theory of the a priori. *Philosophical Perspectives*, 13: 29–55.
—— (2000). A theory of the a priori. *Pacific Philosophical Quarterly*, 81: 1–29.
Cartwright, N. (2007). *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. New York: Cambridge University Press.
Chalmers, D., Manley, D., and Wasserman, R. (eds.) (2009). *Metametaphysics: New Essays on the Foundations of Ontology*. New York: Oxford University Press.
Creath, R. (1991). Every dogma has its day. *Erkenntnis*, 35: 347–89.
Day, T. and Kincaid, H. (1994). Putting inference to the best explanation in its proper place. *Synthese*, 98: 271–95.
De Caro, M. and Macarthur, D. (eds.) (2004a). *Naturalism in Question*. Cambridge, MA: Harvard University Press.
—— and —— (2004b). Introduction: The nature of naturalism. In M. De Caro and D. Macarthur (eds.) *Naturalism in Question*. Cambridge, MA: Harvard University Press.
Dennett, D. (1991). Real patterns. *The Journal of Philosophy*, 88: 27–51.
—— (2003). *Freedom Evolves*. London: Allen Lane.
Deutsch, D. (2011). *The Beginning of Infinity: Explanations That Transform the World*. New York: Penguin Group.
Friedman, M. (1992). *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press.
—— (1999). *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.
——(2001). *The Dynamics of Reason*. Stanford, CA: CSLI Publications.
Glymour, C. (1980). *Theory and Evidence*. Princeton, NJ: Princeton University Press.
Hall, N. (2007). Structural equations and causation. *Philosophical Studies*, 132: 109–36.
Irzik, G. and Grunberg, T. (1995). Carnap and Kuhn: Arch enemies or close allies? *British Journal for the Philosophy of Science*, 46 (3): 285–307.
Kincaid, H. (ed.) (2012a). *Oxford Handbook of the Philosophy of the Social Sciences*. New York: Oxford University Press.
—— (2012b). Mechanisms, causal modeling, and the limitations of traditional multiple regression. In H. Kincaid (ed.) *Oxford Handbook of the Philosophy of the Social Sciences*. New York: Oxford University Press.

Ladyman, J. and Ross, D. (2007). *Every Thing Must Go*. Oxford: Oxford University Press.

Lewis, D. (1994). Humean supervenience debugged. *Mind*, 103: 473–90.

Lowe, E. J. (2002). *Survey of Metaphysics*. Oxford: Oxford University Press.

Maddy, P. (2007). *Second Philosophy*. Oxford: Oxford University Press.

Manley, D. (2009). Introduction: A guided tour of metaphysics. In D. Chalmers, D. Manley, and R. Wasserman (eds.) *Metametaphysics: New Essays on the Foundations of Ontology*. New York: Oxford University Press.

Mantzanvinos, C. (ed.) (2009). *Philosophy of the Social Sciences*. Cambridge: Cambridge University Press.

McDowell, J. (2004). Naturalism in the Philosophy of Mind. In M. De Caro and D. Macarthur (eds.) *Naturalism in Question*. Cambridge, MA: Harvard University Press.

Pearl, J. (2010). *Causality*. Cambridge: Cambridge University Press.

Russell, B. (1959/1912). *Problems of Philosophy*. Oxford: Oxford University Press.

Searle, J. (2009). Language and social ontology. In C. Mantzanvinos (ed.) *Philosophy of the Social Sciences*. Cambridge: Cambridge University Press.

Sober, E. (1988). *Reconstructing the Past: Parsimony, Evolution, and Inference*. Cambridge, MA: MIT Press.

Spirtes, P., Glymour, C. and Scheines, R. (2001). *Causation, Prediction, and Search*. Cambridge, MA: MIT Press.

Williams, M. (2001). *Problems of Knowledge*. New York: Oxford University Press.

Wilson, M. (2006). *Wandering Significance*. Oxford: Oxford University Press.

2

On the Prospects of Naturalized Metaphysics

Anjan Chakravartty

> Ontology and metaphysics, at least in some of their forms, constitute boundary areas into which more strictly scientific analysis gradually shades off, so that there is no sharp line of demarcation between the myths of ontology and the hypothetical entities of empirical science or of mathematics.
>
> Letter from C. G. Hempel to F. Sontag, 14 November 1956, Princeton (Hempel goes on to say that he prefers scientific analysis nonetheless.)

1. A slippery slope

There is a sense, now historically dated, in which believing in many of the unobservable entities and processes described by our best contemporary scientific theories, or the literal (even if only approximate) truth of scientific descriptions of such unobservables, is considered a metaphysical commitment. This is the sense of 'metaphysics' prevalent, for example, in some logical empiricist analyses of scientific knowledge, where claiming literal knowledge of any sort of unobservable is tantamount to metaphysics.

It is fair to say that today, however, epistemic commitments to many canonical scientific unobservables are not generally considered metaphysical per se. There may be good reason, as many scientific antirealists contend, to think that our epistemic powers are not so great as to warrant these commitments, but that is a reflection on epistemology, not the metaphysical nature of the commitments involved. The sorts of beliefs that scientific realists often present as warranted—in unobservable entities

such as positrons and genes, and unobservable processes such as β^+ decay and transcription—are no longer generally viewed as metaphysical in the way that beliefs in properties as tropes or *de re* causal necessity are viewed as metaphysical.

This said, many contemporary philosophers of science and, prominently among them, many scientific realists *do* advocate beliefs concerning things that philosophers today would *still* regard as metaphysical, including beliefs about properties, causation, laws of nature, *de re* modality, and so on. Indeed, philosophical defences of the reasonableness of believing in the sorts of scientific entities and processes that are *not* generally considered metaphysical today, such as genes and gene transcription, often make recourse to views about things that *are* regarded as falling under the purview of metaphysics, such as causation, modality, and so on. Here we have the beginnings of an apparently slippery slope. For if one's account of the reasonableness of believing in gene transcription depends on the reasonableness of one's understanding of the causal relations in virtue of which one is justified in knowing about genes and processes involving them in the first place, it is difficult to see how one could be entitled to the former without the latter. The realist edifice has supports, it seems, in certain metaphysical underpinnings, and the very attempt to establish the integrity of the supports casts one down a slippery slope into deeper and deeper metaphysical theorizing.

I have used scientific realism as an illustration of how slipperiness may come to bear, but it is important to note that in this regard, realism is hardly unique. Arguably, any epistemological position that takes us to have knowledge of the external world—whether of strictly speaking unobservable entities and processes, or of (only) medium-size observable goods, as some antirealist epistemologies would prefer—will face the same challenge. The causal and/or other relations in virtue of which observable things are known by humans themselves act as supports for the reasonableness of our knowledge of the observable, and to furnish a defensible account of these supports is to do what everyone would agree is metaphysics. Here too, the slippery slope seems unavoidable.

Where does it end? It is not my present intention to consider how regresses of philosophical explanation and justification can or must stop. My point here is simply to note that in justifying many beliefs commonly taken for granted, not least in the context of the sciences, the slippery slope presents itself almost immediately. One might well hold, for example, that in order to feel secure in the idea that realism is a coherent epistemic attitude to

take towards the sciences, one should have a defensible account of the processes in virtue of which information is 'transmitted' from the relevant objects of inquiry and scientific instruments, via the array of intermediaries (models, simulations, etc.) we typically employ, to human consciousness. And then, having produced a serviceable account of these processes, one might reasonably wonder whether there is any sense to be made of the notions of property, entity, and so on, that are putatively engaged in them, given the highly variegated natures of things apparently revealed by scientific inquiry. If scientific entities are (*ex hypothesi*) simply collections of properties cohering at locations in space-time, for instance, might it not be reasonable to wonder whether there is a cogent picture of such coherence to be had?

The slippery slope is real, but one may hope nevertheless to keep from falling down. Many philosophers of sciences prefer not to engage in forms of metaphysical theorizing that are very far removed from the ontological theorizing most closely related to scientific inquiry, and this preference comes by way of at least two different motivations. One motivation is pragmatic and reflective of mere differences of philosophical interest: there is, after all, a division of labour in philosophy, and individual philosophers typically focus on the issues that most interest them while others toil elsewhere. The second motivation, however, is at once more principled and less ecumenical, and it is this motivation that interests me here. One might reject any philosophical engagement with the metaphysical underpinnings of various scientific beliefs, because one feels that theorizing this far down the slope is simply too far removed from the details of scientific investigation to be of interest to any interpretation of what scientific theories may say about the world. Or one could go further and suggest that deep metaphysics is too far removed from the details of scientific investigation to yield anything worth having at all. This would be to suggest that engaging in metaphysical pursuits too far down the slope is epistemically impotent, and thus a misguided philosophical pursuit.

Recent philosophy of science has presented both a willingness to grapple with the metaphysical underpinnings of our best current science, as well as a tendency to reject analytic metaphysics as it is commonly pursued in other domains of philosophy. Indeed, some philosophers of science do both,[1] and this suggests at least one strategy for halting the slide

[1] For two recent examples, see Ladyman and Ross (2007), and Maudlin (2007).

down the slippery slope—at least, halting the slide before one proceeds to questions one might think it unprofitable to consider. The idea here is that grappling with the metaphysical underpinnings of our best current science need not amount to metaphysics in the style of analytic metaphysics as it is problematically practiced in other domains. Rather, if done in the right sort of way, metaphysical theorizing might be acceptable in the context of the sciences even if it proves problematic elsewhere. This, in a nutshell, is the proposal and promise of *naturalized metaphysics*.

In the remainder of this chapter, I will examine the thought that if one simply does metaphysics in a naturalized sort of way, one may achieve the twin desiderata of (1) arresting one's slide down the slippery slope in time to avoid philosophizing about matters too far removed from scientific investigation to contribute towards an understanding of the natural world, and thereby (2) avoiding the sorts of metaphysics disparaged by some philosophers of science. I will begin with the assumption that these twin desiderata are, in some form or other, sensible ones to adopt, and argue that the project of naturalized metaphysics has not yet been conceived in such a way as to make these desiderata achievable. As currently described, the very idea of naturalized metaphysics is subject to a debilitating vagueness which renders its advocates unable to articulate convincingly what it is that makes metaphysical theorizing acceptable in some domains and problematic in others.

In section 2, I will explore the idea of naturalized metaphysics in its current form, and consider what it is about some analytic metaphysics that raises the hackles of some philosophers of science. This leads to a worry about the very coherence of naturalized metaphysics, which I consider in section 3. An obvious reply to this worry is presented and revealed to be intuitively compelling but largely empty. In section 4, I begin the process of giving content to this intuitive reply; the content is intended to augment our currently nascent conception of naturalized metaphysics in such a way as to clarify how one might assess different forms of metaphysical inquiry regarding their epistemic potential. To foreshadow my conclusion: it does not appear that this content can provide anything like an 'objective' determination of where one might reasonably stop on the slippery slope, and consequently, there is no 'neutral' advice to be had concerning what sorts of metaphysical theorizing are worth pursuing. The analysis does, however, help us to understand the epistemic risks involved when philosophers engage in different sorts of metaphysical projects. Armed with this

knowledge, an appropriately voluntaristic choice awaits any philosopher who ventures into metaphysics, about where to draw the line.

2. Metaphysics naturalized and *simpliciter*, the a priori, and science

To begin, then, what is naturalized metaphysics? One might think that a clue regarding the nature of naturalized metaphysics should be derivable from the work that many philosophers of science do in investigating the metaphysical implications of our best scientific theories—the job of the metaphysician of science. What ontology of objects and processes is described by the mathematical formalism of theories in fundamental physics? Is natural selection a force that acts on some or other biological entity, or is it simply a statistical outcome of causal interactions acting at other levels of description? There are many questions that arise in thinking about how best to interpret scientific theories that call for some sort of metaphysical analysis, regarding ontology, causation, and so forth. Presumably, then, if there is such a thing as naturalized metaphysics, these sorts of investigations should comprise paradigm instances. However, acknowledging the fact that these investigations all stand in some relation to scientific knowledge, it is far from clear that anything helpful can be learned from simply lumping together such a wide diversity of philosophical projects, and the mere fact that some metaphysics concerns scientific knowledge is hardly an elaboration of the idea that some metaphysics is naturalized.

Let us retreat further, then, to something like first principles, and begin with a working definition of 'metaphysics' *simpliciter*. There is in current philosophical discourse something of a cottage industry whose aim is to determine what metaphysics is, precisely—work in so-called 'metametaphysics'—but it will suffice to proceed in simpler terms here. The Aristotelian conception of metaphysics identifies it principally with two things. The first is the study of being qua being, or ontology: considerations of the most general nature of existence and the natures of things that exist. The presumed generality of this kind of investigation contributes what many have traditionally taken to be a common connotation of metaphysics, that it is an inquiry into universal, timeless truths, but while many metaphysicians aspire to fundamentality of these sorts, it now seems an overly strict constraint in view of much contemporary metaphysics. The second focus of Aristotelian

metaphysics is the study of first causes and theology. Though certain aspects of the ancient study of ontology and first causes are somewhat outmoded today, it is fair to say that a focus on ontology and causation more generally has been retained. For present purposes, then, stripping away some perhaps old-fashioned connotations, let us proceed to think of metaphysics in terms of investigations into ontology and causation.

Having understood metaphysics in these general and innocuous terms, it should be clear immediately that there is nothing here to distinguish metaphysics *simpliciter* from metaphysics pursued in the context of the sciences, since clearly the latter is typified by attempts to theorize about the ontology and causal workings of the various systems and phenomena it investigates, no less than metaphysics *simpliciter*. This of course is what one should expect if naturalized metaphysics is to be a *form* of metaphysics (*simpliciter*), offering an important clue, I believe, in aid of the formulation of a plausible conception of naturalized metaphysics. The distinction between putatively acceptable naturalistic metaphysics and putatively excessive metaphysical inquiry does *not* concern *what* these forms of inquiry aim to do, where the relevant aims are conceived in the general and innocuous terms of shedding light on ontological and causal features of the world. Rather, it concerns *how* these forms of philosophical inquiry go about achieving these aims. It is not in terms of general goals but rather in terms of precise methods that the distinction between naturalized metaphysics and some other brands of ostensibly worrying analytic metaphysics must be drawn.

How, then, is the methodological distinction to be drawn? Metaphysics generally, and to some significant extent, proceeds by way of *a priori* stipulation and theorizing, and produces claims that are empirically untestable. It typically begins with the data of accepted facts and observable phenomena, and then attempts to provide an explanatory account of these things in terms of underlying realities. The degree to which such an account is removed from empirical considerations, however, is highly variable, or so one might reasonably contend. It is the idea of a priori stipulation and theorizing with no significant empirical tethering that generates worries about some approaches to metaphysics relative to others. The a priori character of metaphysics is manifested, in part, in the ways in which its arguments typically proceed, by appeal to intuitions and conceptual analysis. But the untethering of metaphysics from empirical considerations is most profound, so the argument in favour of naturalized

metaphysics would go, in domains of metaphysical theorizing external to the metaphysics of science, where the empirical content of scientific theories and models does not function to restrain otherwise profligate theorizing.

Thus, presumably, the methodological distinction between naturalized and non-naturalized metaphysics is to be understood in terms of proximity to the scientific context. Naturalized metaphysics is metaphysics that is inspired by and constrained by the output of our best science. Non-naturalized metaphysics is metaphysics that is not so inspired or constrained. As a consequence, non-naturalized, untethered metaphysics produces results—theories of universals, substances, bundles, necessities, possible worlds, and so on—that are not conceived as being answerable in any way to empirical investigation. In contrast, naturalized metaphysics, in virtue of its scientific starting point and context, is conceived as being susceptible and sensitive to empirical concerns.

There is a conflation in this first-pass characterization of naturalized metaphysics that should be immediately obvious, and which requires careful unpacking. The suggested distinction between naturalized and non-naturalized metaphysics just rehearsed turns on the idea that the former and not the latter is sufficiently connected, in some way, to empirical findings. This is the force of the idea of constraining otherwise a priori theorizing with empirical data. But in the characterization of naturalized metaphysics just given, it is science that plays the role of constrainer, not empirical data as such. The notion of an empirical constraint is thus conflated here with a scientific constraint. Admittedly, there is a certain caricature of the sciences on which this conflation is benign. The sciences are commonly described as a posteriori investigations whose outputs are empirically testable—observations are made to confirm novel predictions; hypotheses are subjected to experimental testing; instruments and techniques are constructed in order to detect, measure, and manipulate putative entities and processes; and so on. Hence, on this picture of science, the conflation of empirical investigation with scientific investigation in a discussion of naturalized metaphysics may seem entirely reasonable.

This picture of the sciences as comprehensively empirical is nevertheless a caricature, however. That this is so is readily apparent in the fact that not all sciences actually make novel predictions (evolutionary biology), or employ experiments (string theory), or are successful in manipulating things (cosmology). The degree to which and the ways in which the

many domains of investigation that come under the heading of 'the sciences' are empirical is highly variable. As a consequence, the distinction here between a priori and a posteriori methodology cannot simply be superimposed unproblematically on metaphysics and the sciences, respectively. This should raise at least some preliminary concern about the very idea of distinguishing naturalized and non-naturalized metaphysics on the basis of their contact (or lack thereof) with the empirical, simply in virtue of their contact (or lack thereof) with the sciences.

The foregoing note about variability in the empirical natures of different sciences is a telling fact in support of (though ultimately inessential to establishing) an interim contention I wish to put forward now. The contention is that it is a mistake to suggest, as does the preceding caricature, that science is a purely empirical enterprise. The fact that some scientific domains are more highly empirical than others, in more or less impressive ways than others, is some evidence for the proposition that empirical considerations are not exhaustive of what we call science, but even if one were to grant for the sake of argument that all of science is highly empirical in some specified manner, it is still doubtful that it would make any sense to think of the sciences as employing *solely* empirical methodologies. Certainly, throughout most of the history of natural philosophy, what we would now anachronistically identify as scientific investigation clearly incorporated both a priori and a posteriori methods of theorizing, as the line between metaphysics and the new sciences could hardly be drawn with any sharpness.[2] Closer to home, many have argued that there is also good reason to believe that even in the case of the modern sciences, the a priori is inextricably bound up with scientific knowledge.

I will not consider in any detail here the numerous ways in which philosophers of science have documented the central role played by a priori principles and reasoning in modern scientific work, and the corresponding inseparability of the a priori from scientific knowledge produced thereby. In opposition to this contention, one might argue that a priori considerations play no role at all in contemporary scientific practice, which would be to suggest a clean break between the present and clearly established examples of a priori supposition infusing scientific knowledge

[2] See Burtt (1952, on Copernicus, Kepler, Galileo, Descartes, Hobbes, Boyle, and Newton), Buchdahl (1969, on Descartes, Locke, Berkeley, Hume, Leibniz, and Kant), and Woolhouse (1988).

in the past: concepts of substance, essence, and form in the work of Copernicus and Kepler; concepts of the soul and matter for Descartes; concepts of absolute space and time for Newton (recall that his rotating bucket thought experiment is indeed an experiment *in thought*); and so on. If there is one thing that all serious epistemologies of science have in common, however, it is the view that the role played by a priori principles and reasoning in the construction of scientific knowledge is hardly a thing of the past.

One reason for suspicion regarding any suggestion that the a priori has been effectively stripped from the context of the modern sciences is the Kuhnian idea that all periods of normal science incorporate metaphysical assumptions into the disciplinary matrices that make up scientific disciplines and stabilize periods of research. As it happens, this is just one famous example from among the many different accounts to have been elaborated in recent philosophy of science of the cognitive preconditions of scientific work, which aim to describe the prior 'frameworks of investigation', 'networks of concepts', and so forth, that function to establish the very categories of objects, evidence, and inference that allow scientific questions to be posed and then investigated. These frameworks include a priori commitments. On this basis, then, one might plausibly maintain that the very possibility of framing and subsequently probing a hypothesis, empirically, requires that scientists presuppose an ontological scheme of possibilities within which the hypothesis can be formulated, before proceeding according to whatever principles of inference, extrapolation from evidence, and so on, are sanctioned within the relevant scientific community.

There is a rich history of thinking about scientific knowledge in precisely this way. From Kant's emphasis on the conceptual basis of knowledge (not merely grand concepts such as causality, but also finer-grained principles such as those concerning the nature of the universal aether), to neo-Kantian views of what some refer to as the constitutive a priori (Friedman, 2001) or the functional a priori (Stump, 2003). Just as in the case of the metaphysical aspects of Kuhn's disciplinary matrices, these constitutive or functional principles make certain kinds of scientific investigation possible, by providing, inter alia, a conceptual vocabulary and associated definitions in terms of which to cognize reality and fashion scientific ontology. Consider, furthermore, conventionalist understandings of the geometry of space-time, other variants of neo-Kantianism such as internal realism (Putnam, 1981), the 'scenes of inquiry' (Jardine, 1991),

'styles of reasoning' (Hacking, 1992), and even more broadly, Foucauldian epistemes. The contention that the sciences incorporate a priori commitments as part of their modus operandi—as a prerequisite to doing scientific work and thus generating scientific knowledge—is hardly controversial.[3] Indeed, it has been a widespread contention for some time now:

> A scientific theory arises in response to a problem . . . for instance one of producing a consistent explanation capable of accounting for both the wave-like and particle-like aspects in the behaviour of light. But a problem . . . presupposes a relatively stable matrix—a reality scheme, an intelligibility scheme, a *Lebenswelt*, basically, a conceptual matrix sufficiently consistent so that problems can arise within it. . . . Scientific theories deal with problems which arise within an intelligible context. Proposals establishing such a context, defining a reality-matrix, are not scientific theories but *metaphysical proposals*. (Kohak, 1974: 24, in an early paper on the demarcation of physics from metaphysics.)

In fairness, it is important to note that at least some neo-Kantian conventionalists claim, or are presented as claiming, that by adopting such views, they are not really engaging in metaphysics as such. In the present context, however, this apparently conflicting diagnosis is revealed as mere terminological confusion following from a particular use of the term 'metaphysics'. Some neo-Kantian positions (internal realism is an exemplar here) do reject a particular conception of metaphysical knowledge in terms of what is sometimes called 'metaphysical realism', but even this leaves the door open to neo-Kantian metaphysics. And since the very point of Kant's Copernican revolution was to *fuse* metaphysics and epistemology in such a way as to transcend the scepticism he saw as an inevitable consequence of the empiricism and rationalism that preceded him, it would be a mistake to think that the opposition of some neo-Kantians to metaphysics targets metaphysics *simpliciter*, as I have elaborated it here. Rather, it targets a particular metaphysical assumption.

However it is conceived very precisely, we are now in a position to ask an important question regarding the a priori: what is the import of the role it plays in scientific work, and of the resultant sculpting of scientific knowledge in accordance with various a priori moulds sanctioned within different domains of scientific investigation? In the next section, I will

[3] In Chakravartty (2010), I argue that all systematic epistemologies of science appeal in some way, shape, or form to the notion of a priori content, sometimes explicitly and otherwise implicitly.

argue that the prevalence of a priori content described by the epistemologies of science just mentioned leads to a difficulty for the project of naturalized metaphysics as it is currently conceived. The solution to this difficulty will require an elaboration of the project in more compelling terms than have been articulated by its proponents thus far.

3. The incoherence of naïvely naturalized metaphysics, and a reply

Let us proceed with the understanding that scientific knowledge harbours some a priori content. I believe that this generates a charge of incoherence against the idea of naturalized metaphysics. And while an intuitive response to this charge is easily furnished, the response reveals just how impoverished our current conception of naturalized metaphysics is, or so I will suggest. I will consider these assertions in some detail momentarily, after first clarifying two implicit and critical assumptions: the first concerning the nature of scientific progress; and the second concerning the version of naturalism at stake in this discussion.

There is, I believe, in the very conception of naturalized metaphysics as it is currently best conceived, an implicit, non-trivial assumption about the nature of scientific progress. Any attempt to do metaphysics on the back of the sciences might be viewed in at least two different ways. On the one hand, there is a well-established tradition in the history of philosophy that regards metaphysical theorizing as a fruitful heuristic for scientific work. This is to suggest that to the extent that a priori theorizing is part of scientific work, it does not take the form of constitutive or functional principles per se, but rather speculative possibilities that are ultimately converted into empirical propositions as scientists devise means of testing them empirically. Popper, for example, maintained that metaphysical theorizing is heuristically useful to science in allowing the development of concepts that ultimately suggest directions for empirical research programmes. Atoms, elements, corpuscles of light, and fluidic or flow-type views of electricity are all examples of concepts that were born metaphysically, he would say, but grew up empirically. The same is true of now disconfirmed ideas such as the existence of a luminiferous aether, or traceable particle identity over time, or universally deterministic laws. These notions

may also have begun as a priori commitments, but over time, scientific work has revealed them to be untenable, empirically.[4]

Consider a more recent example from medical research in the fields of cell and tissue biology.[5] A significant literature here suggests that two conflicting metaphysical presuppositions—'reductionism' and 'organicism'—have shaped studies of cancer over the past few decades. Reductionism in this domain takes the form of genetic determinism, the idea that certain biological states and processes can be explained wholly or primarily in terms of genes, and organicism is the view that emergent phenomena at higher levels of biological organization are crucial to these explanations. Some organicists, for example, believe that the production of cancer cells can be explained in terms of abnormal tissue organization. But here, just as in the case of Popper's examples, one might argue that to the extent that the a priori infuses the relevant scientific knowledge, it is merely in the form of hypotheses that are offered in anticipation of a posteriori investigation. That is to say, the a priori is merely a heuristic device.

If naturalized metaphysics is to constitute a different and better form of metaphysics than some alternative, however, the merely heuristic conception of a priori theorizing cannot exhaust what is intended. The notion that the a priori may serve as a potential expedient in the service of empirical research seems uncontroversial, but there is nothing in the *merely* heuristic conception to explain how one might distinguish between naturalized metaphysics and any other sort, except in retrospect. That is to say, one might look back over the history of the sciences and describe metaphysical theorizing that was ultimately fruitful of empirical research as naturalized metaphysics post hoc. But surely this is not (exhaustive of) what anyone hoping to do naturalized metaphysics can intend, since the relevant intention is fuelled by the aspiration to distinguish naturalized from non-naturalized metaphysical theorizing *in the present*—not least as a normative guide to what sorts of metaphysical projects should be considered most worth engaging now. This is not to say, of course, that a heuristic role for the a priori is incompatible with the idea of naturalized metaphysics, but rather that this cannot be the whole story.

[4] Not everyone holds the influence of a priori theorizing in the context of scientific work to be so benign. Duhem, for instance, believed that metaphysics is often counterproductive to science, because it sometimes opposes or attempts to subordinate promising empirical investigations.

[5] For a detailed discussion of this case, see Marcum (2005).

In order to add to this story in the manner required, an assumption about scientific progress is necessary: one must assume that some parts of scientific theories are likely to be retained over time across theory change, and furthermore, that we are in a position to identify at least some of these parts. Without some such identification as a basis for metaphysics, the scientific ground of naturalized metaphysics would inevitably shift significantly in time, raising serious doubts about the motivation for distinguishing between metaphysics bearing a privileged relation to empirical science and metaphysics that lacks this quality. For if the scientific basis were radically unstable, one would have no good reason to suspect that metaphysics done in conjunction with it at any given time is preferable to metaphysics that is alien to it. The assumption of stability in the progress of science is by no means trivial, but certainly, many epistemologies of science are compatible with it. Some forms of instrumentalism and empiricism, for example, are compatible with the notion that the empirical content of scientific theorizing survives changes in theoretical superstructures, and several forms of scientific realism suggest that there are criteria according to which one may identify aspects of scientific theories that are likely to survive theory change in mature domains of science.[6]

A second assumption necessary to the articulation of the concept of naturalized metaphysics concerns the form of naturalism at issue. The idea of naturalism is generally associated with two rather different theses. The first we have encountered already: the notion that some philosophical (in this case, metaphysical) questions and answers evolve into and rightfully become scientific-empirical questions and answers over time, as thinking about them matures. This is what Quine suggested, for example, regarding epistemology and empirical psychology, or regarding natural kind philosophy and scientific taxonomy. This cannot be what naturalized metaphysicians have in mind, however, for as we have just noted, their enterprise is much diminished if we have no reason to think that this way of doing metaphysics is capable of telling us something about the world now, as opposed to merely spinning its wheels on the off chance and wishful thinking that what is produced may evolve into empirical investigations that tell us something about the world later. The force of this point is augmented by the observation that much of the metaphysics of science

[6] For some recent accounts of 'selective' realism, see Chakravartty (2007a), French (2006), Ladyman (1998), Psillos (1999), and Worrall (1989).

concerns issues—the nature of causation, laws, and modality; the objective or subjective basis of natural taxonomy; the individuality or non-individuality of entities in fundamental physics—that we have no reason at all to suspect will *ever* be settled by empirical investigation alone; on the contrary, since the concepts involved are not wholly empirical, or for that matter, wholly scientific.

A second and rather distinct idea commonly associated with naturalism is that philosophy (in this case, again, metaphysics) is continuous with science, and it is this conception of naturalism that is relevant to the present discussion. Presumably, the naturalized metaphysician holds that metaphysical theorizing in a naturalistic vein is continuous with and thereby close to the ground of empirical results (recall the quotation from Hempel with which this chapter began), unlike other work in metaphysics that is clearly further away. Here we see the impetus for the naturalized metaphysician's rejection of what I earlier described as metaphysical theorizing that is too far down the slippery slope from scientific investigation to be of serious interest to an interpretation of scientific knowledge, and the even more severe rejection of such theorizing as a misguided endeavour: epistemically impotent with regard to its prospects for yielding any genuine understanding of the natural world. Continuity with science, then, is the suggested means by which to ensure that metaphysics does not lapse into the unprofitable excesses of non-naturalized metaphysics. Naturalized metaphysics, *ex hypothesi*, in virtue of its continuity with science, enjoys some degree of epistemic privilege.

Having elaborated what I identified at the start of this section as two critical assumptions, concerning the nature of scientific progress and the form of naturalism at issue here, let us now proceed to the charge of incoherence against naturalized metaphysics. The metaphor of continuity is highly suggestive, but how should it be cashed out more precisely? Interestingly, our current conception of naturalized metaphysics, as characterized by those who are sympathetic to it, does not generally advance beyond the provision of further, equally vague sentiments, though even this modicum of help is instructive. It is not uncommon to hear that continuity in this context is evidenced by the fact that naturalized metaphysics is 'derived from', 'based on', or otherwise 'inspired' or 'motivated' or 'constrained by' our best science, which thereby serves as the proper

'ground' for metaphysical theorizing.[7] The fact that these expressions in quotation marks are vague, and the implication that we may wish to say something more precise about the notion and epistemic value of continuity, will be the subjects of section 4. For the moment, however, let us work with what we have, and consider more carefully the family of relations constituted here by derivation, basing, inspiration, motivation, constraint, and grounding.

What is it about continuity with the sciences, and the putatively resultant grounding of naturalized metaphysics, that is meant to afford it a privileged status in comparison to non-naturalized metaphysics? Recall that it is the a posteriori, empirical content of the sciences that is supposed to enhance the credentials of the metaphysics of science, and thus, by extension, one might argue that metaphysics that is done in such a way as to be empirically grounded may claim some epistemic privilege. Here we have the proposal of a criterion of adequacy for bona fide metaphysics: it must stand in a certain kind of *sanctioning* relation to empirical inquiry. And herein lies the difficulty. In order to make sense of the idea that one body of belief—scientific knowledge—stands in a sanctioning relation to another body of belief—the results of some metaphysical theorizing—it must be possible to distinguish clearly the relata of the sanctioning relation one from the other. That is, one must be capable of clearly distinguishing the associated forms of inquiry, so that one can then ground the outputs of one on the outputs of the other. But note here, once again, that there is a conflation in this reasoning! It is one thing to entertain the idea of grounding a priori theorizing in a posteriori knowledge, but it is quite another to imagine grounding a priori theorizing in scientific knowledge. For as we have seen, scientific knowledge itself has a priori dimensions.

Thus we arrive, finally, at the worry that there is something apparently incoherent about naturalized metaphysics as it is currently conceived. On this conception, metaphysical theorizing is legitimate only insofar as its constraint or ground is something empirical. However, in practice, this legitimization is attempted by taking *science* to be the constraint on or ground of proper metaphysical theorizing, and the sciences by their nature cannot provide the purely a posteriori content that the suggested criterion of legitimacy requires, because scientific knowledge comprises a blend of a

[7] Ladyman and Ross (2007, pp. 37–8) offer a more specific take on continuity when they recommend the use of metaphysical theorizing in the service of scientific unification.

priori and a posteriori content. While metaphysics *simpliciter* may be described purely in terms of its a priori or non-empirical character, the sciences cannot, on pain of caricature, be described purely in terms of their supposed a posteriori or empirical character. Given that scientific knowledge does have a priori dimensions, the sanctioning relation proposed by naturalized metaphysics, between a priori and a posteriori content, simply cannot be realized in the manner suggested. As a consequence, the idea of naturalized metaphysics as it is currently conceived lapses into incoherence.

There is an obvious reply to this charge of incoherence. Granting that scientific knowledge has a priori dimensions, and even granting that different branches of the sciences are a posteriori in highly variegated ways and to highly variable degrees, it remains the case that the forms of inquiry we collect under the banner of the sciences are permeated with a posteriori content in virtue of the empirical concepts with which they are concerned. So why not take 'naturalized metaphysics' to label those metaphysical projects that are derived from, based on, inspired by, motivated by, constrained by, and grounded in this specifically empirical content, as opposed to scientific knowledge more generally? Given that most scientific inquiry is inescapably infused with empirical content in virtue of a posteriori investigation, one might seek to ground naturalized metaphysics in this same, specifically empirical content. What could be simpler? I take this obvious reply to the worry of incoherence to be intuitively compelling—there seems something right about the idea that what distinguishes some forms of metaphysical theorizing from others is the question of how closely connected (or not) these projects are from the specifically empirical content of scientific knowledge.

Unfortunately, however, and despite its intuitive appeal, this natural way of thinking about naturalized metaphysics is largely empty, and cannot do the work our intuitions might suggest it should. The problem is that the criterion of legitimacy suggested is far too easy to fulfil. Indeed, there is good reason to think that it is generally trivially satisfied, which would entail that *every* metaphysical project is an instance of naturalized metaphysics: clearly a poor result from the perspective of an aspiring naturalized metaphysician. To illustrate the point, consider the following tendentious example. There was a philosopher who maintained that the theory of the Forms and realism about universals is derived from experience. For in the course of making empirical observations, he noticed that

various objects of his experience had a number of similarities and differences, and these observations were borne out in countless numbers of a posteriori investigations. He then theorized about what ontological features these objects of empirical study might have in order to account for all of the observed similarities and differences, and *voilà*, Platonism is derived from experience.

It is hard to imagine anyone being especially impressed by the *empirical* nature of this derivation, though it is fair to say that it followed entirely in the course of theorizing based on empirical observations. This is an extreme demonstration of the emptiness of attempting to explicate the idea of naturalized metaphysics in terms of vaguely specified linkages to empirical content and a posteriori investigation. Any metaphysical project that is not immediately self-undermining *ab initio* will be consistent with empirical observations, and thus too easily linked to empirical content if the terms of the linkage are specified too broadly. In the next section, I will begin the process of articulating a more robust conception of naturalized metaphysics, by taking seriously the notion that the terms of this articulation must be significantly more precise than those we have canvassed thus far.

4. The grounding metaphor

We have learned that the idea of naturalized metaphysics must go beyond the mere idea of metaphysics as a useful heuristic for scientific work, and that its distinctive character, in contrast to non-naturalized metaphysics, has something to do with its continuity with a posteriori, empirical investigation. The ways in which this continuity is conceived as facilitating a distinctive character for naturalized metaphysics, however, have not yet been spelled out in a way that secures the distinction. There is a strong will here to distinguish cases in which, though it may never be possible to carry out an empirical test—for example, to establish the one-way speed of light, or to detect the presence of hidden variables in quantum mechanics—it is nevertheless possible to understand what may be regarded as a priori commitments as appropriately linked to a posteriori content. This linking takes the form of some appropriate grounding in a system of empirical concepts, observations, and so on. What is required, then, is some means by which to distinguish such work from work in non-naturalized metaphysics, which is perceived as being too preoccupied with epicycles on issues whose consideration takes place a very long distance from empirical investigation.

Recall that the explicit goal of naturalized metaphysics is to tame the putative excesses of some metaphysics *simpliciter* (which lends itself, or so the assertion goes, to inappropriately constrained speculation about things one could not possibly hope to know about, or about things concerning which one may have no reason to suppose there are even facts of the matter) by linking metaphysics to the naturalistically respectable project of trying to interpret the empirical content of our best science. This linkage must be to empirical content and not merely to science, for as argued in section 2, it is a mistake to think that scientific knowledge is wholly empirical, and as argued in section 3, the project of naturalized metaphysics must focus on empirical content if it is to escape the charge of incoherence. Escaping this charge, one might then hope to avoid falling down the slippery slope into the darkest depths of metaphysical speculation. In order to achieve these goals, we must scrutinize more seriously the metaphor of grounding, and as a first step, I suggest that we pay closer attention to the idea of *continuity* with the empirical. What does continuity mean, in this context? This, I believe, is the crux of the issue, and the only hope for giving a defensible formulation of the idea of naturalized metaphysics. In pursuit of greater clarity here, we require some sort of metric or metrics by which to make more precise the relevant notion or notions of continuity. Armed with such metrics, the more precise meanings of expressions like 'closeness to' and 'distance from' empirical work may finally come into focus.

The project I suggest here is a large one, and I cannot claim to know all or how many such metrics may be relevant to explicating a fully compelling account of naturalized metaphysics. One must begin somewhere, however, and in the remainder of this chapter, I will describe two parameters that appear to play a central role in our thinking about metaphors of proximity and distance with respect to a posteriori investigation. Even in so incomplete a form, I believe that these reflections yield important morals for our assessment of the epistemic potency of much of what is typically identified in contemporary philosophy as the metaphysics of science. And as we shall see, even our best efforts to demarcate naturalized metaphysics may leave open the question of where precisely to dig in one's heels on the slippery slope.

Perhaps the most obvious way to think about proximity to empirical content is in terms of what I will call the 'experiential distance' of an object of inquiry. This concerns the manner in which it is detected, if in fact it is detectable at all. Tyson, the barking dog across the street in our otherwise

quiet neighbourhood, is directly detectable by me using my unaided senses. Proteins are less directly detectable; I would need to take a sample from Tyson and perform an assay in the lab to detect them. The possible worlds in which I now demonstrate this procedure to my friends and neighbours are not detectable at all. There is a spectrum of cases here, and the further one moves along the spectrum, the further the distance of the object of inquiry from perception by the unaided senses. Since the demise of thoroughgoing rationalism in the philosophy of science, it is widely held that the further one goes in terms of experiential distance, *ceteris paribus*, the weaker the epistemic power of our inferences concerning putative targets of investigation. Of course, this is not to say that experiential distance is strictly inversely correlated with inferential strength, since the relevant epistemic conditions are not always equal. Neither is it to say that weaker inferences are insufficient to produce knowledge; scientific realists of various stripes argue precisely this point—when it comes to certain unobservable things under certain conditions, one may have good reasons to infer their existence. The idea is rather simply that the epistemic challenge to make warranted inferences mounts with experiential distance, *ceteris paribus*.

Another way of understanding the notion of proximity to empirical investigation is in terms of what I will call 'risk', which concerns how susceptible a hypotheses or a theory is to disconfirmation in light of the results of empirical work. The idea of 'susceptibility' here is a measure of how strongly empirical evidence weighs on our assessments of truth and falsity. For example, if empirical considerations are judged to be relatively unimportant to the assessment of the truth value of a proposition, the risk it engenders is low. Hypotheses that make very precise novel predictions about observable phenomena, on the other hand, take a greater risk than those that do not, which may include, for example, hypotheses that merely accommodate data that is already known. Hypotheses that include epicycles and idle wheels in accommodating the same empirical data as those that do not take no extra risk, and may thus be judged negatively as a result. Hypotheses that are riskier in these and perhaps other senses are generally viewed as being closer to empirical investigation, *ceteris paribus*, and the closer to empirical investigation they are, the greater the confirmation boost they receive if their predictions and accommodations are borne out in empirical investigation. As in the example of experiential distance, the idea of risk also generates a spectrum of cases, and one's assessments of

these factors help to determine the degrees of belief one should associate with the relevant hypotheses and theories.[8]

Experiential distance and risk are two parameters that seem central to cashing out the metaphor of grounding in a way that clarifies how continuity with a posteriori investigation can be ordered by means of epistemic metrics.[9] With tools such as these, expressions like 'closeness to' and 'distance from' empirical work begin to take on more precise meanings, which then give substance to our conception of naturalized metaphysics. These are merely two among what may well be a collection of parameters that are relevant to this conception, however. The onus is on those of us who see promise in the idea of naturalized metaphysics to work out precisely what further factors one might justifiably consider in determining when a given piece of a priori theorizing can be grounded in a posteriori investigation in such a way as to meet a reasonable epistemic threshold or standard, to whatever extent is deemed appropriate in the context. The ultimate promise and defence of naturalized metaphysics awaits this yet further articulation.

I am optimistic about the prospects of naturalized metaphysics, both as a philosophical endeavour that can be distinguished from metaphysics *simpliciter*, and as a form of inquiry that may contribute to our knowledge of the world. Lest optimism lead inadvertently to dogmatism, however, let me conclude with a brief, cautionary observation for aspiring naturalized metaphysicians (such as myself). If there is a tendency among some analytic metaphysicians to ignore the outputs of empirical science at their peril, there is equally a tendency among some aspiring naturalized metaphysicians to court excessive confidence in their wealth of scientific knowledge, but this latter vice is no less philosophically counterproductive than the former.

Quite reasonably, one may ask: where do the speculations of metaphysicians of science typically fall on the spectra of experiential distance and risk described above? Undeniably, on reflection, it seems the answer to this question must be: not very close to empirical investigation! Typically, no

[8] An interesting topic that I will not pursue here is the relationship between experiential distance and risk. One might expect these parameters to be sometimes but not always inversely correlated: sometimes, because hypotheses and theories about entities and processes at greater experiential distances are sometimes less susceptible to empirical findings; but not always, as for example when they generate novel predictions.

[9] For further consideration of these parameters, see Chakravartty (2007b).

matter how well informed by empirical details emanating from scientific work, our accounts of fundamental ontology in the sciences, or laws of nature, or the nature of time, and so on, do not fare especially well with respect to the metrics of experiential distance and risk. What this reflection serves to highlight is the perhaps obvious fact that expressions like 'derived from', 'based on', 'motivated by', 'inspired by', 'constrained by', and 'grounded in' do not mean *entailed by*. Indeed, from an epistemic perspective, the best one can hope for in employing these expressions is that metaphysics that stands in such relations is *compatible with* our best empirical science, and this should prove an antidote to any danger of hubris on the part of aspiring naturalized metaphysicians (such as myself). Despite too much casual rhetoric to the contrary, the metaphysical theses argued for by metaphysicians of science are not extracted from the empirical content of science, as if they were there already simply waiting to be mined. They are developed by means of a priori theorizing in the course of interpreting scientific claims. A necessary condition for successful interpretation here is compatibility with the science at issue, but this condition can only take one so far.

How impressed should one be with *mere* compatibility? Recall the tendentious example of realism about universals. Mere compatibility does not buy one much epistemic warrant in the absence of an impressive assessment of values of parameters such as experiential distance and risk. Test cases from the metaphysics of sciences abound, and what is striking about these cases is that the relevant a priori theorizing is highly underdetermined by our best science. Our best physics, for example, does not determine the fact that reality ultimately consists in an ontology of fundamental relations lacking ontologically significant relata (cf. Ladyman and Ross, 2007), or that laws of nature should be taken as primitive, that there are no such things as universals, and that time passes (cf. Maudlin, 2007). These are not the sorts of things on which current physics can be expected to pronounce univocally one way or the other. Indeed, aspiring naturalized metaphysicians often go out of their way to promote metaphysical theses that play no role in scientific practice (for example, pertaining to the unity of the sciences, or the metaphysical status of laws of nature), or that are at odds with at least some scientific claims taken at face value (such as the routine reference to ontologically significant relata, or the apparent inconsistency of Special Relativity with the passage of time).

Of course, it is an epistemically rational strategy to make inferences that are informed by the best information one has available, and on the reasonable assumption that mature scientific claims produced in part by empirical investigation likely furnish better information than their negations, metaphysics that is compatible with the outputs of our best science is obviously preferable.[10] The point here, however, is a deeper one: without invoking criteria such as experiential distance and risk, there is no obvious reason to think that metaphysics that is 'derived from' our best science is any more likely to produce knowledge of the world than metaphysics that is not so 'derived', but nonetheless compatible with our best science. And since our best empirical evidence, scientific or otherwise, generally underdetermines our best metaphysical theorizing, there is something deeply confused about any proposal for naturalized metaphysics that would seek to save metaphysics simply by scientizing it. It is for this reason that I contend that the attempt to sanction some metaphysics by making the grounding relation the *sine qua non* of legitimate metaphysics does not get us anywhere worth being all by itself. If we are to be naturalized metaphysicians, let us dedicate ourselves to the philosophical analysis of epistemically probative metrics for assessments of grounding.

That we already employ such metrics in the philosophy of science today, if only implicitly, explains why it is that much of what was once considered metaphysical by some empiricists is no longer considered metaphysical by most. Theorizing about the nature of things like phlogiston and white blood cells seems rather close to the ground of empirical investigation. Typical theorizing in the metaphysics of science, including some of the examples I have mentioned here, are not close at all. Let us be entirely transparent, then, about what naturalized metaphysics can achieve. It brings a priori considerations to bear on developing accounts of the underlying features of entities and processes of scientific discourse. It applies familiar criteria: consistency; coherence; simplicity; scope; unification; minimizing ad hoc hypotheses and primitive concepts; and so on. It marshals intuitions about these desiderata, in just the same manner as non-naturalized metaphysics, and appeals to these intuitions in determining the relative strengths and weaknesses of metaphysical hypotheses and

[10] For probing scepticism regarding this assumption, however, see Monton (2011), which argues that compatibility with current physics is not a plausibly truth-conducive desideratum for metaphysics.

theories. It also adduces intuitions regarding which phenomena most require explanations, and what would count as a good one. Generally, these arguments have the form, either explicitly or implicitly, of inferences to the best explanation. The better the explanatory work a hypothesis does, by some proffered lights, the greater its warrant. If the experiential distance is great and risk small, however, this should be reflected in our degrees of belief.

Metaphysical inferences will never be as strong as we might like, even if they are naturalized, and given the nature of inference to the best explanation, there will always be ultimately irresolvable subjective differences in our assignments of degrees of belief based on largely irresolvable differences in some of the intuitions we bring to bear on their assessment. As a consequence, it is simply a mistake to think that there is any one place on the slippery slope that is an objectively rational place to stop, for determinations of where best to stop are inevitably subject to variable intuitions regarding how much experiential distance and risk is tolerable in an inquiry we engage in hopes of learning something about the world. These determinations are choices that all metaphysicians must make, even those who pay close attention to the sciences. Some philosophers, perhaps many, will see this as bad news for metaphysicians of science. It is difficult to see how it could be, however. This is merely the human epistemic condition.

Acknowledgements

Aspects of this chapter have benefited greatly from discussions with a number of audiences: at the annual meetings of the Eastern Division American Philosophical Association, British Society for the Philosophy of Science, and Dubrovnik Philosophy of Science Conference; at colloquia hosted by the Universities of Cologne, Toronto, and Victoria, and McMaster University; and at a meeting of the Bay Area Philosophy of Science group. For their support and insight, I am especially grateful to Paul Humphreys, Don Ross, Paul Teller, and Bas van Fraassen.

References

Buchdahl, G. (1969). *Metaphysics and the Philosophy of Science: The Classical Origins, Descartes to Kant*. Oxford: Blackwell.

Burtt, E. A. (1952). *The Metaphysical Foundations of Modern Science*. Garden City: Doubleday.

Chakravartty, A. (2007a). *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.

——(2007b). Six degrees of speculation: Metaphysics in empirical contexts. In B. Monton (ed.) *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas C. van Fraassen*, pp. 183–208. Oxford: Oxford University Press.

——(2010). Metaphysics between the sciences and philosophies of science. In P. D. Magnus and J. Busch (eds.) *New Waves in Philosophy of Science*, pp. 59–77. London: Palgrave Macmillan.

French, S. R. D. (2006). Structure as a weapon of the realist. *Proceedings of the Aristotelian Society*, 106: 167–85.

Friedman, M. (2001). *Dynamics of Reason*. Chicago: University of Chicago Press.

Hacking, I. (1992). *'Style'* for historians and philosophers. *Studies in History and Philosophy of Science*, 23: 1–20.

Jardine, N. (1991). *The Scenes of Inquiry: On the Reality of Questions in the Sciences*. Oxford: Oxford University Press.

Kohak, E. V. (1974). Physics, meta-physics, and metaphysics: A modest demarcation proposal. *Metaphilosophy*, 5: 18–35.

Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science*, 29: 409–24.

——and Ross, D. (2007). *Everything Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Marcum, J. A. (2005). Metaphysical presuppositions and scientific practices: Reductionism and organicism in cancer research. *International Studies in the Philosophy of Science*, 19: 31–45.

Maudlin, T. (2007). *The Metaphysics Within Physics*. Oxford: Oxford University Press.

Monton, B. (2011). Prolegomena to any future physics-based metaphysics. In J. Kvanvig, ed., *Oxford Studies in Philosophy of Religion, vol. III*, pp. 142–65. Oxford: Oxford University Press.

Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth*. London: Routledge.

Putnam, H. (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.

Stump, D. J. (2003). Defending conventions as functionally *a priori* knowledge. *Philosophy of Science*, 70: 1149–60.

Woolhouse, R. S. (ed.) (1988). *Metaphysics and Philosophy of Science in the Seventeenth and Eighteenth Centuries: Essays in Honour of Gerd Buchdahl*, The University of Western Ontario Series in Philosophy of Science, vol. 43. Dordrecht: Kluwer.

Worrall, J. (1989). Structural realism: The best of both worlds?' *Dialectica*, 43: 99–124.

3

Scientific Ontology and Speculative Ontology

Paul Humphreys

A conflict has emerged in contemporary philosophy between two quite different ways of approaching ontology. Over the last few years a growing divide has emerged between the fields that are often called 'analytic metaphysics' and 'scientific metaphysics.' Analytic metaphysics is characterized by the importance given to a priori methods and to conceptual analysis. A posteriori scientific results are at best a peripheral part of its evidence base. In contrast, scientific metaphysics looks to the results of contemporary science for guidance to the correct ontology.[1]

This tension between scientific and analytic metaphysics has a long history. The logical empiricists tried to do away with metaphysics, failed, and metaphysics made a spectacular comeback. Its revival increased the potential for friction between the two areas, although the antipathy seems to be more visible to philosophers of science than it is to metaphysicians. Some philosophers of science have principled objections to metaphysics or to conceptual analysis, often for reasons stemming from a sympathy for empiricism or for naturalism.[2] Others object to the use of non-perceptual intuitions as a reliable method for arriving at truths.[3] Still others find the

[1] Although scientific metaphysics tends to draw its conclusions from fundamental physics, I do not presuppose the correctness of that approach here. One reason is the failure of various reductionist programs and the empirical evidence for emergent properties that are not fundamental in the sense that they fall outside the domain of high energy physics, digital physics, information-theoretic physics, or whatever is currently considered to be 'fundamental.'

[2] See e.g. van Fraassen (2004) for an example of the first orientation.

[3] See e.g. Williamson (2005).

pretensions of analytic metaphysics objectionable.[4] My paper is an attempt to lay out some of the reasons behind the tension and to suggest the form that appropriate ontological inquiries might take. There is a legitimate place for metaphysical reasoning and conceptual analysis in the philosophy of science. It is also true that parts of contemporary metaphysics are indefensible.[5]

Put in its simplest terms, the question is this: If one has reason to think that broadly aimed anti-metaphysical arguments have failed and that metaphysics can be pursued as a supplement to science, what methods should a scientifically informed metaphysics employ? As a starting point, I shall assume that some version of realism is acceptable, where by 'realism' I mean any ontological position that accepts the existence of at least some entities such that those entities and their intrinsic and relational properties do not depend on conscious intellectual activity.[6] This allows that some entities are so dependent, such as the property of fear, that some are culturally dependent, such as the existence of money, and that our own access to them may be mediated, perhaps unavoidably, by representational apparatus. The realism assumption is not unassailable, but there are few idealists nowadays, neo-Kantians can interpret the arguments given here as establishing choice criteria for the representational and evidential apparatus to be employed, and it allows many entities of traditional metaphysics, such as abstract objects, causal relations, haecceities, and others to be considered under the definition of realism.

This general characterization of realism is important for two reasons. First, more specific kinds of realism are defined in opposition to some other position, such as empiricism in the case of scientific realism, nominalism in the case of realism about universals, or constructivism in the case of realism about mathematical objects. One cannot ignore those oppositions entirely because they have to a large extent shaped the tradition of what counts as metaphysics, and it is the lack of clear success for the anti-realist member of each of those pairs that has left open the possibility for the

[4] See Ladyman and Ross (2007).

[5] In this article I shall discuss only selected examples of mainstream analytic metaphysics. There are more extreme forms but they are easier targets of criticism. In some cases the defense of a religious position appears to be the ultimate goal of the arguments.

[6] One thing that is not allowed to exist: conscious intellectual activity by a deity. This will be a restriction only if the existence of such a deity is used to transform some version of what is ordinarily considered to be idealism into a deity-dependent form of realism.

broader kind of realism considered here. But those oppositions are not directly relevant to my concerns. For example, the differences between scientific metaphysics and analytic metaphysics need not be due to differences about the status of abstract entities, as the older opposition between realism and nominalism suggested. Scientific metaphysics has to consider the possibility of abstract entities because first order logic is too weak to serve as the basis for mathematics that is adequate for science and at least some second-order mathematical properties such as 'is a random variable' are, if not purely formal, abstract.[7] Secondly, the kind of broad ontological categories that emerge from these oppositions—abstract versus concrete entities, observable versus non-observable entities, and so on—are too crude to be methodologically useful. The kinds of procedures that are relevant to establishing the existence of a set with the cardinality of the real numbers are radically different from those that would be needed to establish the existence of non-natural moral properties.

Although the tension between the philosophy of science and analytic metaphysics is prima facie about what exists, that is proxy for a deep difference about what counts as an appropriate philosophical method. This means that the epistemology of ontological claims will play a central role in the discussion. I thus have a second assumption. It is that the methods of the physical sciences have been successful in discovering the existence and properties of a number of generally accepted entities in the physical sciences.[8] These procedures include empirical, mathematical, and computational methods. This assumption should also be reasonably uncontroversial and it does not deny that other methods can and have been successful when applied to appropriate subject matter. The assumption is incompatible with versions of universal scepticism under which all methods of inquiry are equally unable to provide warrants for truth but on those views the basic dispute under discussion here has no content. Weaker positions such that specifiable scientific methods are less fallible than the methods of analytic metaphysics could be substituted for our second assumption.

[7] The nominalist program for scientific mathematics pursued by Field (1980), while revealing, is not widely regarded as successful.

[8] By 'physical sciences' I mean astrophysics (including cosmology and cosmogony), physics, and chemistry. I would be willing to extend this to some, but not all, biological sciences but this is not necessary for present purposes.

The two assumptions together provide us with an overall strategy. Realism suggests that because things are the way they are independently of how we know (of) them, and there is no prima facie reason to hold that a single, uniform set of epistemological methods will be successful in all domains, we should begin with the position that different domains of inquiry might require different epistemological approaches. Our second assumption suggests that we begin with those methods that have evidence in favor of their success in a given area and explore how far they can be extended before they break down and need supplementation. This suggestion does not preclude a priori methods and appeals to intuition being successful in areas such as mathematics and metaphysics. It relies on the inductive fact—and it is a fact—that those methods are unreliable when applied to the domain of the sciences. That is, we already have evidence that such methods fail when they are applied to the sciences mentioned above unless they are supplemented by empirical data. The project now is to see if incursions into metaphysics by scientific methods can be successful. It would be naive to simply assume that methods which have been successful in the physical sciences will also be successful for ontological purposes in other areas, but a good deal of sound epistemic practice is enshrined in the methods of the more rigorous sciences and it can be useful to see whether they can be adapted to philosophical applications. If and when these methods break down, their failures can provide clues as to what should replace them in the given domain.

Having set up the broad opposition in these terms, I shall immediately restrict the scope of my discussion. It is impossible to precisely capture the scope of analytic metaphysics, which allows some practitioners to deny that they contribute to the discipline. So rather than discussing the broader category of 'analytic metaphysics,' I shall consider what I call 'speculative ontology.'[9] Consider ontology to be the study of what exists, a domain that includes not only what there is and what there is not, but the intrinsic

[9] The term 'speculative metaphysics' is often used to designate a certain type of activity, including the kind of metaphysics to which Kant was opposed. I shall therefore avoid that term. The term 'naive ontology' is not quite right because the lack of appeal to scientific knowledge can be willful, although the effects are the same as would result from ignorance. The term 'speculative mathematics' was traditionally used to denote what is now known as pure mathematics (see Stewart et al., 1835), a tradition not to be confused with how speculative mathematics is now conceived (see Jaffe and Quinn, 1993).

properties of and relational structures between the entities, including second-order and higher properties and relations.

What restrictions does science place upon ontological results? I do not have a sharp criterion to distinguish scientific results from other kinds of conclusions. One reason is that considerable parts of contemporary cosmogony are so remote from direct empirical tests that they cannot be assessed primarily on grounds of empirical adequacy and are, in their own way, as speculative as are parts of metaphysics. But such activities are constrained by well established scientific results such as general relativistic accounts of gravity, conservation principles derived from high energy physics, and the Standard Model, together with widely accepted standards of scientific evidence such as consistency with empirical knowledge, use of the principle of total evidence, explicitness of the assumptions used, willingness to abandon a hypothesis in the face of counter-evidence, and so on. All of these constraints are defeasible but they must be respected by those wishing to take into account all of the relevant contemporary evidence. The relevant differences between scientific ontology and other kinds of ontologies thus tend to center on what are considered constraints on the truth of the basic assumptions and the evidence base and when the kinds of constraints just mentioned are applied, we have scientific ontology. When they are not, we have speculative ontology.

Scientific ontology goes beyond whatever ontology the current scientific community currently endorses because underdetermination considerations often suggest the exploration of alternative ontologies that are compatible with scientific principles and evidence and in which scientists display little interest.[10] Two examples are Bohm's interpretation of classical quantum mechanics (Bohm, 1952) and David Chalmers' conclusions about dualism (Chalmers, 1997). These are serious attempts to argue for an alternative to the mainstream ontology and both satisfy the conditions for a scientific ontology. I shall now present four principal reasons why speculative ontology should be viewed with deep suspicion. These are its widespread factual falsity, its appeal to intuitions, its informal use of conceptual analysis, and its assumption of scale invariance. These objections are not independent but because they do not apply uniformly to all activities within speculative ontology it is worth considering them separately.

[10] Thus the deferential attitude to science endorsed in the natural ontological attitude (Fine, 1984) is inadequate for an informed scientific ontology. I shall say more on this point in section 7.

1. The factual falsity of speculative ontology

Some claims of speculative ontology can be dismissed on the straightforward grounds that they are not even true, let alone necessarily true. The fault is not that these claims are false. Falsity is a ground for criticizing a claim, but not for dismissing it as irresponsible. What moves a claim from the first to the second category is the existence of an epistemic situation consisting of these components: i) the claim is incompatible with the results of a well-established knowledge base, ii) those results are widely known in the relevant epistemic community, and iii) a modest amount of work would lead to those results being understood by philosophers. A murkier realm looms when the fact that the claim is false is brought to the attention of those who argue for it and the fact is ignored.

Traditional metaphysics tended to concern itself with necessary truths and although they were truths of our world too, the claims were generally beyond the reach of empirical science and so any potential conflict with scientific results was muted. Two things have changed that situation. The increased sophistication of the physical sciences has led to a number of metaphysical theses that were widely held to be necessarily true being shown to be factually false by experimental evidence. Well known historical examples are the identity of indiscernibles, universal determinism, and universal causation.

Conversely, there has been a move by some metaphysicians towards taking certain metaphysical claims as contingent, rather than necessary. A much-discussed contemporary example of this kind of project is David Lewis's Humean supervenience program, the physicalist position within which the ontology consists only of local points of space-time and local properties instantiated at those points. Everything else globally supervenes on that. Here is what Lewis had to say about the modal status of his view: 'I have conceded that Humean supervenience is a contingent, therefore an empirical, issue. Then why should I, as a philosopher rather than a physics fan, care about it? . . . Really, what I uphold is not so much the truth of Humean supervenience as the *tenability* of it. If physics itself were to teach me that it is false, I wouldn't grieve' (Lewis, 1986: xi).

As an exercise in theorizing, Lewis's attitude would be unobjectionable were the position seriously put to an empirical test. But it has not in the sense that science long ago showed that Humean supervenience is factually false. The claim that the physical world has a form that fits the constraints of Humean supervenience is incompatible with well-known and well-

confirmed theoretical knowledge about the non-separability of fermions. It was empirically established before Lewis's position was developed that entangled states in quantum mechanics exist and do not supervene on what would in classical cases be called the states of the components. This feature of our world is sufficiently well confirmed as to make Humean supervenience untenable.[11]

It would be unreasonable to object to an intellectual activity simply on the grounds that its products did not represent truths about the concrete world. Many areas of mathematics have their own intrinsic pleasures and there is virtue in understanding how the world might have been, but is not. Yet when reading many Lewisian disciples, they do little to dispel the impression that the apparatus under construction is a guide to the ontological structure of our world and one wonders why so much effort is expended on this apparatus when more promising ontological alternatives are waiting to be investigated.[12]

Speculative ontologists might object to this line of reasoning on the grounds that what I have called factual falsity is better described as an inconsistency between a sentence that is supported by empirical evidence and a sentence supported by a priori reasoning, that the truth value of the scientific sentence is, on inductive historical grounds, uncertain and liable to eventually be rejected and hence that we should not allow the empirical evidence for the scientific claim to take priority over the a priori reasons for the other claim.[13] We can reply to this concern in two ways. First, that worries stemming from the pessimistic meta-induction have been exaggerated; that raising the mere logical possibility that a scientific claim might be false adds nothing to what we already knew, which is that such claims are contingent; and that in the absence of specific evidence that is directly relevant to a particular scientific claim which suggests that the claim is false and the presence of specific evidence that it is true, the rational epistemic action is to accept the claim as true. Such worries are more pressing when claims about fundamental ontology are involved and this response does indeed gain purchase there. But many conflicts between scientific and speculative ontology are not about fundamentals.

[11] See e.g. Maudlin (2007), Chapter 2.

[12] For a defence of Lewis's tenability claim, see Weatherson (2010).

[13] I am grateful to Anjan Chakravartty for raising this objection.

Despite initial appearances, Lewis' metaphysics is as much about preserving common sense as it is about establishing esoterica.[14]

The second reply is to accept the response and to note that what science can do is to provide us with descriptions of possible situations that, even if they are not true of our world, expand our set of known metaphysical possibilities and can also show that certain metaphysical generalizations are not necessarily true. It is often as important to be aware of error as it is to know the truth and science is capable of informing us about both, within the limits of fallibility that are inescapable in both science and ontology.

2. Whose intuitions?

A second source of discontent with speculative ontology is its appeal to intuitions as evidence. Although this suspicion is well-founded in many cases, arguments against the use of intuitions are not straightforward. Empirical evidence has shown that intuitions are highly fallible about both factual matters and reasoning, violating our principle of least epistemic risk (see section 6), but that evidence usually comes from studies on agents with skills below the expert level and we should be wary of too hasty conclusions about the inappropriateness of all appeals to intuition.[15] I take an intuition to be an unreflective and non-inferential judgment and I shall be concerned primarily, although not exclusively, with non-perceptual intuitions.[16] It is common to add that intuitions are immediate, but insisting on immediacy qua temporal immediacy is unnecessary, for what is important is the unreflective aspect. Were I to be asked for my judgment on some issue, then immediately fell into a prolonged state of unconsciousness within which no subconscious or unconscious reasoning processes occurred, and I finally

[14] For a defense of this claim, see Symons (2008).

[15] In mathematics, views are divided about the use of intuitions. There is a stark contrast between the position of the logician J. Barkley Rosser: 'The average mathematician should not forget that intuition is the final authority' (quoted in De Millo, Lipton, and Perlis, 1979) and that of Frege: 'To prevent anything intuitive from penetrating [the realm of arithmetical reasoning] unnoticed, I had to bend every effort to keep the chain of inferences free from gaps' (Frege, *Begriffsschrift* in van Heijernoort, 1967).

[16] One use of the term 'intuition' ruled out by this definition is its use in the expression 'Here's the intuition', often said by mathematicians and other practitioners of formal methods after presenting some technical item. What is being conveyed is an informal understanding, stripped of necessary but secondary details, and that understanding is the result of significant reflection on the content of the formal apparatus.

rendered my judgment upon awakening, or the intervening time was spent mentally rerunning the last six miles of a recent marathon, my response would be as much an intuition as any instantaneous judgment. There are other uses of the term 'intuition' in contemporary philosophy. In addition to the sense just given, it is common to use the term to refer to a psychological entity, a propositional attitude accompanied by a felt sense of certainty.[17] Brief reflection on the proposition involved is also allowed. Sometimes an intuition can be the starting point for a process of reflective equilibrium, but more often it is a non-negotiable item.

Two obvious problems with intuitions are that they tend to differ, often considerably, between philosophers and they are often wrong when they are not about everyday experiences because we have no prior knowledge base on which to draw. These are familiar objections but I want to draw attention to a further reason to doubt many appeals to intuition. In traditions in which logical analysis and explicit definitions are central, it is often taken for granted that conceptual analysis and the use of intuition are subject matter independent activities and are largely agent-independent as befits their foundational status. But there are good reasons to doubt both of these claims and to hold that the reliability of intuitions and the use of conceptual analysis is domain specific and varies between agents. That is, a given agent's intuitions can be a reliable source of knowledge when applied to one domain and unreliable when applied to another while another agent's intuitions have the inverse degrees of reliability.[18]

This claim is clearly true for perceptual intuitions. Certain qualia are epistemically accessible by the visual modality but not others, such as flavours which must be accessed by the faculty of human taste. Other domains, such as those of mathematics and of modal truths, are assumed to

[17] For this use, see Bealer (1999: 247). Bealer argues that one can have an intuition that A and not believe that A because certain mathematical falsehoods can seem to be true yet one knows that they are not. This is arguable—if you do not believe A, you usually do not have the full force of a traditional intuition. It is not enough just to 'seem' that A is true otherwise a guess will count as an intuition. I doubt that there is a uniform answer here, in part because different philosophers use the term 'intuition' differently. Note that the modifier 'rational' can be applied to intuitions, for example when extrinsic constraints such as consistency with other intuitions or Bayesian coherence criteria are imposed. For further discussions of intuitions, see DePaul and Ramsey (1998).

[18] Reliability is also agent-specific in that different agents have different degrees of reliability in a given domain, a fact that is often ignored in the exchange of intuitions in philosophical arguments.

be accessible by rational intuition or spatial intuition but are beyond the reach of olfactory intuitions.[19] Different agents tend to have different degrees of acuity, and so an idealization is made to ideally competent agents. In the perceptual realm, most humans were considered to have capacities sufficiently close to the ideal standard to enable them to come to know most observational truths. This was also thought to be true in the realm of a priori truths for ideally rational agents, and in that case the vast majority of humans were considered able, if not willing, to attain the standards needed for simple cases of a priori knowledge. Yet this level of generality is implausible.

Let us separate two claims. The first is that for a given agent, appeals to a priori intuitions have equal validity across all domains in which non-empirical knowledge might be available. This is demonstrably false. Intuitions about probability are unreliable for most people, one famous example being the Monty Hall problem.[20] There is a correct answer to the problem but most people arrive at a wrong answer when they approach the problem intuitively and their intuitions must be corrected either by arguments or by real or simulated frequency data.[21] In contrast, intuitions about the transitivity of preferences or the correctness of modus ponens are, outside a few problem cases, generally correct.[22] The second claim is that most agents are sufficiently close to ideally competent across all domains when using intuitions. This is also demonstrably false. When considering geometrical intuitions, number theoretic intuitions, probabilistic intuitions, moral intuitions, logical intuitions, and so on, a given individual often has strengths in one area and not in another. So just as there are different perceptual modalities for different observable properties, and different scientific instruments for different detectable properties, the evidence seems to suggest that there are different levels of success for

[19] This surely is not necessarily true—there could be an arithmetic of smells—but I know of no mathematician who has reported working in this way.

[20] For an extensive discussion of this problem see Rosenhouse (2009).

[21] Paul Erdös was initially certain that the accepted solution was incorrect and was convinced only after seeing the results of a computer simulation.

[22] Anyone who has taught logic to enough students will have encountered the curious phenomenon of an occasional student who seems to lack the ability to understand some primitive logical inferences. Sometimes they function using concrete inferences; in other cases one suspects that associations between categorical statements substitute for hypothetical reasoning, for example substituting the sequence 'A customer complained. Call the supervisor' for inferences based on the rule 'If a customer complains, call the supervisor.'

different kinds of intuitions, and that we must specify a domain of reliability for a given type of intuition and a given agent.

Of course, there have been attempts to reduce expertise in these disparate domains to a common basis. But in cases in which intuitions involve concepts, the burden of proof is on those who claim that all mature agents have equal reliability when appealing to intuitions, regardless of the concepts involved. This is at odds with what we know from practice. Some philosophers have better moral intuitions or logical intuitions than others, where the basis for evaluating reliability is long-term consistency between an individual's intuitive judgments and evidence gained from non-intuitive sources. This conforms to the widely held view in science and mathematics that physical, biological, mathematical, and other types of intuition come in degrees and there is no obvious reason why this should be different in philosophy. Furthermore, because in scientific areas the expert 'intuitions' are usually, although not invariantly, shaped by considerable knowledge of the relevant science, the non-inferential characteristic of intuitions will hold only at the level of conscious processing and this will also be characteristic of philosophical intuitions.

Behind the tacit assumption that the use of intuition is the basis of a general method often lies an appeal, inherited from the tradition of linguistic analysis, to competent users of a language. But semantic externalism and the division of linguistic labor entail that this competence, and the use of linguistic intuitions in conceptual analysis, has severe limitations.[23] If the semantic content of some part of language depends upon the way the world is, then knowledge of the world, often quite detailed knowledge of the world, is required to know what constitutes that content. Perhaps this is a minor difficulty since the use of intuitions is often restricted to core examples of the correct use of the expression.[24] Even so, there is a related and more serious problem that occurs independently of whether one subscribes to externalism and it again requires us to decide whose intuitions should prevail. Consider the extension of the predicate 'dangerous'. Applied to other humans, many of us have some reasonably reliable intuitions about whether a given individual has the associated property or not and can identify some core examples of

[23] Arguments for semantic externalism can be found in Putnam (1974).

[24] Although not always because considerable philosophical discussion occurs about borderline cases of concepts.

dangerous humans. Yet we would, or should, defer to those who have better intuitions, such as members of the fire department, war veterans, or members of mountain rescue teams when circumstances demand it. What are the analogous circumstances in ontology and who are those with superior intuitions? I place little credence in appeals to intuitions and so I pose this challenge to the ontologists who use them. How do you train philosophers to improve their intuitions and how can we recognize when you have been successful? Who granted the Doctor of Intuition diploma on the metaphysician's mental wall?

We have arrived at the following principle: All methods have their domains of successful application and associated probabilities of error. There is not a uniform method of 'appeal to intuition' that has the same success rate for all rational agents across all domains. In the presence of conflicts between intuitions and in the absence of criteria establishing who are the expert practitioners in a given domain, we should remain agnostic about any appeals to intuition and reject them as a source of evidence.

3. Conceptual analysis

Conceptual analysis has a legitimate role to play in scientific ontology. Providing explicit or implicit definitions for concepts is an essential part of scientific understanding and should be encouraged but concepts drawn from everyday experience are often the wrong ones to analyze, whether in scientific or metaphysical contexts. One reason is that everyday, psychologically grounded, concepts usually do not have sharp, necessary and sufficient conditions, so that when an explicit definition is set up, it will not fit all intuitive examples of the concept. This is often the basis for the kind of unproductive exchange well-known to those who attend philosophy colloquia. In scientific contexts, conceptual analysis is complicated by the use of approximations and idealizations and a great deal of care is needed to draw the distinction between an idealized theoretical concept and the concept that is used in applications. Despite these reasons for maintaining a clear distinction between refined, theoretically based, concepts and concepts drawn from less considered sources, appeals to intuition have begun to serve as a replacement for traditional conceptual analysis in some areas of philosophy and even sophisticated philosophers with high standards for argumentation have suggested that their use is acceptable.

Take Frank Jackson's characterization of conceptual analysis: 'conceptual analysis is the very business of addressing when and whether a story told in one vocabulary is made true by one told in some allegedly more fundamental vocabulary' (Jackson, 2000: 28). This characterization covers not only familiar cases that lie in the vicinity of analytic truths and falsehoods, but the kinds of traditional reductive procedures in the philosophy of science that considered whether condensed matter physics is simply re-describing complex combinations of phenomena for which high energy physics has its own language.[25] Presupposed in this account is the view that we can tell when we are describing the same things in different ways. When this decision procedure relies solely on our understanding of some language, appeals to a priori knowledge are reasonably unproblematical. But the situation is different in discussions of reduction because the higher-level concepts are often presented to us through a theory at the higher level and not through explicit definitions or a quasi-reductive compositional account of the ontology. The identification of the referents of the different descriptions then often requires empirical knowledge and cannot be carried out a priori, a fact that is reflected in necessity of identity claims resting on the contingent truth of the identity. Furthermore, the evidence for the existence of the higher-level phenomenon is empirical, not a priori. For example, purely theoretical, *ab initio*, derivations of high temperature superconductivity are infeasible and the approximations used to arrive at the higher-level theory are justified on experimental grounds, not on a priori reasoning and definitions.

Not all contemporary conceptual analysis takes the traditional form of definitions. For example, David Chalmers and Frank Jackson have argued that conceptual analysis based on explicit or implicit definitions is unavailable and unnecessary in many epistemological contexts. 'It is sometimes claimed that for $A \supset B$ to be a priori, the terms in B must be definable using the terms of A. On this view, a priori entailment requires definitions, or explicit conceptual analyses: that is, finite expressions in the relevant language that are a priori equivalent to the original terms, yielding

[25] The question of whether the Nagelian bridge laws used in that tradition are empirical or definitional is obviously relevant to whether conceptual analysis in this sense is sufficient for this example. The fact that neither the concepts of condensed matter physics nor its ontology are reducible to those of high energy physics is but one reason why many mereological claims fall into the category of speculative ontology.

counterexample-free analyses of those terms. This is not our view. . . . If anything, the moral of the Gettier discussion is that . . . explicit analyses are themselves dependent on a priori intuitions concerning specific cases, or equivalently, on a priori intuitions about certain conditionals. The Gettier literature shows repeatedly that explicit analyses are hostage to specific counterexamples, where these counterexamples involve a priori intuitions about hypothetical cases' (Chalmers and Jackson, 2001: 320–2). On this view, intuition plays a central role but since far more than a mastery of language is needed to bridge the levels in realistic reductionist examples, this position is less than convincing.

Goldman (2007: 18) argues that 'Philosophical analysis is mainly interested in common concepts, ones that underpin our folk metaphysics, our folk epistemology, our folk ethics, and so forth.' That would be fine, and indeed it may be the best approach for many issues in ethics, for ethics cannot move too far from what seems instinctively correct to most people without encountering overwhelming resistance. The situation is different in ontology because metaphysical arguments have led to highly counter-intuitive positions. The fact of being counter-intuitive is not itself an objection, for science frequently uses counter-intuitive representations too, but these are usually supported by empirical data. The problems arise when metaphysicians venture judgments about domains for which the appropriate concepts are far removed from, and often alien to, our human intuitions. Humans are epistemic experts in certain limited domains of inquiry, those that are best suited to our naturally evolved cognitive apparatus, and this includes many ordinary moral judgments. Had we evolved in a different environment, it is likely that our concepts would have been different, probably very different. Bees are the product of evolution, but there is little reason to think that they have similar concepts to us.[26] This suggests that when considering ontological issues, we should also consider cognitive agents who are not subject to our own contingent limitations and are different from or superior to us in cognitive abilities.

An example from computational science may illustrate why this appeal to extended cognitive abilities is useful in certain situations. We are delegating significant parts of computationally based science to non-human executors

[26] There is empirical evidence that bees perceive the world differently from us and thus would be likely to have a bee ontology that is different from human ontology. See Chittka and Walker (2006).

in computational neuroscience, complexity theory, condensed matter physics, and in many other domains, a trend that is destined to accelerate. We shall then with increasing frequency require novel conceptual frameworks to first, employ those techniques in the most effective possible ways and secondly, to understand the results emerging from their deployment. This is because what makes an effective representation for a computer is frequently different from what we humans find transparent in the conceptual realm.

Consider how computerized tomography represents data. An X-ray beam is directed through an object and impinges on a detector which measures the intensity of the received X-rays. If f(x,y) represents the intensity of the X-rays at the point <x,y>, then the line integral along the direction of the beam L, which is represented by the Radon transform of f along L, sums the intensities at all points along the line. Rotating the source around the object and plotting the values of the Radon transform for each of these angles gives a sinogram, which is what the detectors 'see'. It is rarely possible to know what the irradiated object is by visually inspecting the sinogram. Yet computers, using the inverse Radon transform, can easily produce data that, when turned into a graphic, 'directly' represent the object, such as a human skull, in a form accessible to the human visual system.[27]

The sinogram image has no compact representation in human languages connecting it with the ordinary perceptual concept of a human skull and since the machines operate on purely extensional representations, this example directly raises the problem that intensional characterizations of properties in natural languages are likely to be too sparse to capture the kinds of extensionally characterized properties used by computational science. What constitutes a pattern for a human is more restricted than (or is just different from) what constitutes a pattern for a computer. This suggests that philosophers need to do one of three things; expand their repertoire of representations to include those used by computational devices and instruments, despite the fact that those representations do not conform to any that are intuitively accessible to humans; concede that certain current and future scientific activities lie beyond human understanding, assuming that understanding requires possessing the relevant intensional representations, a position that is at odds with the fact that

[27] For examples of sinograms see <http://demonstrations.wolfram.com/ComputedTomographySimulationUsingTheRadonTransform/> accessed 23 May 2012.

we can understand at least some of these activities; or push their explorations of sub-conceptual, non-representational, and non-conceptual content beyond the realms currently explored in artificial intelligence, such as the use of dynamical systems theory or neural net representations. The second and third options entail that there are parts of ontology that lie beyond what is currently intuitively accessible, while the first option is unlikely to be successful without a considerable amount of conceptual retraining.

4. Philosophical idealizations

Consider the following claim. Having insisted that a collective, long term cognitive effort by humans allows an approximation to the ideal cognitive conditions required for a strong modal tie of intuitions to the truth, the author goes on to write that:

> Some people might accept that the strong modal tie thesis about intuition . . . [is] non-empirical but hold that [it does] nothing to clarify the relation between science and philosophy as practiced by human beings. After all, these theses yield only the possibility of autonomous, authoritative philosophical knowledge on the part of creatures whose cognitive conditions are suitably good. What could this possibly have to do with the question of the relation between science and philosophy as actually practiced by us?
>
> The answer is this: The investigation of the key concepts—intuition, evidence, concept possession—establish the possibility of autonomous, authoritative philosophical knowledge on the part of creatures in those ideal cognitive conditions. The same concepts, however, are essential to characterizing our own psychological and epistemic situation (and indeed those of any epistemic agent). (Bealer, 1998: 202–3)

So what are these idealized cognitive conditions? In the article by Chalmers and Jackson cited in section 3, the authors write: 'A priority concerns what is knowable in principle, not in practice and in assessing a priority, we idealize away from contingent cognitive limitations concerning memory, attention, reasoning, and the like' (Chalmers and Jackson, 2001: 334). But this view quickly leads to idealizations that are questionable and commit us to epistemic abilities that can trivialize philosophical conclusions. In the realm of idealizing human abilities, if we were allowed to sufficiently idealize our limited abilities, an appeal to supernatural agents would always be available. Moreover, since it has been remarked that the correct metaphysics will be

revealed at the limit of idealized epistemology, some, although not all, epistemological idealizations of the 'in principle' kind will render certain claims of analytic metaphysics unassailable by default.

Although these approaches often go by the name of 'in principle' approaches, they are more accurately seen as forms of idealization and this requires us to answer the question: What counts as a legitimate philosophical idealization? The analogous question for scientific domains has been much discussed, but there is little in the way of sharp answers to the philosophical question.[28] Just as it is appropriate to require of scientific idealizations that a given idealization maintains some contact with properties of real systems, we should similarly require that we have some criteria for how to relax philosophical idealizations to bring them into contact with human abilities. One kind of epistemic idealization involves the extrapolation of human abilities. We can perform arithmetical operations on integers of a certain size and we then extrapolate that ability to much larger integers that human memory capacity is incapable of storing. That kind of idealization seems to be philosophically legitimate for some purposes because we know the algorithm for generating integers, we are familiar with memorizing small collections of them, and extrapolating that familiar experience to larger collections is just 'more of the same'.[29] A different kind of epistemic extension involves augmentation. Examples of this are easy to give for sensory modalities. We are not naturally equipped to detect the spins of elementary particles and we require instruments the output of which must be converted into a form that we are equipped to detect. Here it is unreasonable to suggest that detecting spins involves an in-principle idealization of human abilities, for there is nothing like that ability in our perceptual repertoire.

The epistemology of speculative ontology relies heavily on cognitive extrapolations and augmentations; the appeal to the perfections of God in the ontological argument, but against the standard theological position that, lacking the full complement of those perfections themselves, it is blasphemous for humans to suggest that they can fully understand the concept of God; the appeal to truths at the limit of scientific inquiry, when

[28] The literature on the differences between those who appeal to ideally rational agents and those who require bounded rationality is relevant here.

[29] This argument does not apply when we are accounting for how much of science is applied in practice.

we have no conception of what limit science will be like; oracles, or Platonists, who can inexplicably access mathematical truths, and so on.[30] Whether these are legitimate extrapolations or augmentations of human intellectual powers is an open question in the absence of criteria for acceptable philosophical idealizations.

5. Scale variance

Here is another principle: *Whether the world is scale-invariant is a contingent fact*. There is evidence that it is in certain ways and that it is not in others. Humans are all inescapably middle-sized objects with limited cognitive capacities that developed in the course of dealing with properties associated with similarly sized objects. One of the great scientific shifts in perspective was the twentieth-century realization that physical phenomena were not universally scale-invariant, that classical mechanics did not extend unchanged to the realms of the very small and the very large (nor, as a result, was it exactly correct in the realm of the middle-sized).[31] Once we had arrived at that realization, the status of the human a priori and of human experiences as sources of knowledge should have changed dramatically. Yet much work in epistemology and metaphysics proceeds as if that scientific shift had never occurred.

What properties, relations, objects, laws, and other parts of our ontology are like at scales radically different from those with which we are natively equipped to deal cannot be inferred from the evidence of direct experience. Consider this simplistic argument: ordinary experience tells us that a magnet has two opposite poles. If one cuts a bar magnet in two, each half has both a north and south pole. Therefore, by induction on common experience, one concludes that magnetic monopoles should not exist. And indeed, magnetic monopoles have not to this point been detected. But the scientific arguments for and against the existence of magnetic monopoles are of a very different kind than the simplistic argument just given and

[30] For clarification: the ontology of limit science here is included in the category of speculative ontology because there is no currently available way of inductively inferring from the present state of science to its limit state.

[31] I exclude here renormalization theories that focus on scale-invariant phenomena. Ladyman and Ross (2007) have discussed similar issues under the idea of 'scale-relative ontology'. Their discussion of Dennett (1991) on which their scale-relative position is based is very helpful in understanding the consequences of Dennett's paper.

current versions of quantum field theory suggest the existence of dyons, of which magnetic monopoles are a special case. For a less naive inference, take the limit of zero resistance in a suitable conducting material (usually achieved by lowering the temperature) and you have moved into a distinctively different physical realm, one in which not just Ohm's Law is false but the ontology changes as the conductive carriers, which originally are electrons, are replaced by a Cooper pair superfluid.

So what morals should we draw from the lack of scale invariance? First, the distinctions between the observable and the unobservable, between the non-theoretical and the theoretical, and between entities at the human scale and those outside that scale are different distinctions. The fact that the first two draw on different categories, objects, and properties in the first case, terms in the second, has often been noted. The logical independence of the elements of the second and the third distinctions can be shown by the fact that money is a classically human scale construct but there exist multiple different theoretical measures of what falls under its scope—M0, M1, M2, and M3 are common measures. Also, the description 'a sphere twice the size of the largest sphere considered to be at the human scale' is non-theoretical, whatever non-trivial sense of 'theoretical' you use, but describes an entity beyond the human scale. The other two cases are easy. The fact that the third and the first distinctions are incommensurable becomes evident if we focus on properties rather than objects. Properties can be observable or not, but size scales do not apply to properties, including metric properties such as 'is one foot long,' only to regions of space possessing them.

Now consider the standard definition that induction involves inferences from the observed to the unobserved. Stock examples of inductive inference, such as generalizations from sub-populations to the whole population, which involve claims that the same property possessed by members of the sub-population will be possessed by all members of the population, can be misleading because usually all instances of the generalization exist at the same scale. But inferences from the observed to the unobserved across differences of scale introduce an additional inductive risk. In addition to the usual assumptions such as that the future resembles the past and that properties are projectable, we must make the assumption that the part of the world under investigation is scale-invariant. This kind of inference is often classified as extrapolation, and extrapolation is, at root, a particular kind of inductive inference and is hence not a priori. Thus what might seem to be a result of a priori reasoning within speculative ontology,

and this is especially true of various forms of mereology when put forward as a comprehensive ontological position, actually contains a hidden a posteriori element that is, moreover, likely to be false. This argument applies independently of what particular perceptual abilities an agent has, as long as they are not universal, and the concepts involved in the ontology require empirical access to be acquired.

Much of metaphysical argumentation has deep similarities with the arguments that accompany thought experiments and when constructing a hypothetical scenario, we need to know the laws and specific facts that apply at the scale at which the scenario is described. The only source of that knowledge is science; our intuitions and a priori imagination and reasoning are not equipped to provide it. At least some of the problematical examples of speculative ontology arise from simply assuming scale invariance when the system does not exhibit it. For example, the telegraphic identity 'water = H_2O' is false as stated since the left hand side refers to the usual macroscopic fluid and so the right hand side should read 'a macroscopic collection of H_2O molecules interacting in such a way that the properties of liquidity, transparency, ability to undergo phase transitions, and so forth are present.'[32]

Here then is a central problem of using intuitions in speculative ontology. The apparently a priori methods mask a tacit appeal to an inductive inference when making the scale-invariance assumption. Because human intuitions about ontology are obtained by experience with human-sized entities, any inference to regions beyond that domain involves an inference from the empirically known to the empirically unknown and that contains an inductive risk. Because there is considerable evidence that this inductive risk is high and that the conclusions of similar inferences have turned out to be false, generalization from intuitions in speculative ontology should be avoided.

6. The principle of least epistemic risk

One feature that traditional metaphysics and traditional empiricism had in common was a commitment to risk aversion. It can be captured in this principle: When competing ontological claims are made, determine the

[32] For a related point, see Johnson (1997).

degree of epistemic risk associated with the methods used to establish or deny the existence of the entity in question and make the ontological choice based on the method with the lowest risk. The risk need not be zero—certainty may not be obtainable—but in their different ways, the foundational enterprises embedded in traditional metaphysics and traditional empiricism both relied on this principle, together with a second principle that inferences from the foundations never decrease the degree of risk, making the foundations the most secure of all claims.[33]

This epistemic principle has important consequences for ontological claims. For example, it makes the opposition to scientific realism by traditional empiricism inappropriate because traditional empiricism takes observables to constitute a fixed category tied to human perceptual abilities and it takes beliefs about observables to be uniformly less risky than beliefs based on what are considered to be unobservables.[34] I criticized the first assumption in Humphreys (2004), Chapter 2; here I focus on the second assumption. Suppose we take the human perceptual apparatus as the foundation for all beliefs about concrete entities. In this role, the human perceptual apparatus plays the same function as calibration devices in the sciences. If we believe that a room measured with a laser rangefinder has a length of 22 feet, this is based on the user's belief that were he to use a tape measure instead, he would directly see the coincidence between the 22 feet mark on the tape and the wall of the room. Assessments of systematic and random errors for an empirical quantity are always made with respect to a calibration standard for the 'true' value. But the device used for the calibration standard is rarely the one used to collect data in most experiments because the conditions under which data from a particular system must be collected are not those under which the calibrating system has optimal functionality. To illustrate the different domains of applicability of traditional empiricism and scientific empiricism, at the time of its painting a portrait by Vermeer could have been authenticated by his wife who witnessed him painting it but now, X-ray fluorescence and wavelet decomposition techniques would supercede any human judgment. Differently, although the correctness of individual steps in a long

[33] Of course inferences together with additional evidence might result in a decrease.

[34] I have couched the discussion here in terms of beliefs to accommodate traditional empiricist positions although my own view is that this anthropocentric apparatus should be abandoned.

computation can be determined a priori by humans, the balancing of foreign exchange trading by Deutsche Bank AG is made by automated methods, and unavoidably so. The philosophical moral is this: the fact that there exist agents who act as a ground for beliefs about a given epistemic task does not entail that those agents can act as a foundation for comparative judgments about risk for that task under all conditions. As science gains greater access to what exists, the judgments of humans, whether empiricists or speculative ontologists, become increasingly risky and so will be replaced by conclusions drawn from scientific sources.

7. Why philosophers of science cannot avoid doing metaphysics

Some of the problems of speculative ontology have been described. Does this mean that we should avoid ontological or metaphysical arguments entirely? This is not possible and to see why consider the following question. How did we get to the point where it is necessary to once again eliminate certain types of metaphysical reasoning? Here is one broad historical thread that gives at least part of the reason. By its very nature, much of metaphysical activity involves decisions about issues that are under-determined by science. In particular, for decisions about ontology, this route was signaled by Carnap's characterization of external questions as having conventionally chosen answers and by Quine's relativized ontology. (See Carnap, 1956 and Quine, 1969.) Wanting such choices to be based on less vague grounds than the pragmatic criteria offered by those two philosophers, some metaphysicians, understandably, have tried to provide convincing arguments for the choice of one ontology over another.

This is a legitimate role for philosophy. Ordinary scientific practice is not oriented towards establishing claims of realism or anti-realism. For the most part, when scientists try to do philosophy, they do it as amateurs, with noticeably poor results.[35] Science is designed to obtain results from complex experimental apparatuses, to run effective computer simulations,

[35] There are exceptions. Richard Feynman, despite his often caustic comments about the philosophy of science, had a number of important philosophical insights, some contained in Feynman (1967). Other scientists, equally critical of the philosophy of science, are simply naive, which is what one would expect of those with no training in the field. For one example, see Hawking and Mlodinow (2010).

to derive predictions from models, along with other tasks unrelated to ontology. In arguing for his Natural Ontological Attitude in which one leaves the choice of ontology to the relevant scientific community, Arthur Fine suggested, in analogy with Hilbert's programme in metamathematics, that 'to argue for realism one must employ methods more stringent than those in ordinary scientific practice' (Fine, 1984: 85–6). This analogy with consistency proofs is misplaced. The justification procedures involved in using observational or experimental techniques to assert the existence of some entity or type of entity are not stronger than the theoretical representations involved in making the realist claims, merely different.[36] Moreover, even in formal sciences, philosophical arguments about existence are relevant. Here is an example. Using a consistency proof, Herzberg (2010) has shown that infinite regresses of probabilistic justification do not render justification of the end-state proposition impossible but he notes that the value of the justificatory probability may not be expressible in the form of a closed form solution. Because there is no precise, universally accepted, definition of a closed form solution, reasons for and against the acceptability of probability values that cannot be estimated are not mathematical but recognizably philosophical, whether presented by philosophers or by mathematicians.

Some scientific activities, such as the search for the Higgs boson or the earlier search for the conjectured planet Vulcan, also have the aim of establishing the existence or non-existence of certain features of the world and philosophers can learn much from the procedures used. There is a strong tradition in arguments for and against realism that criteria for what exists must be general. Yet one moral that can be drawn from claims of existence made within science is that the evidence for existence claims is often domain-specific. This suggests that a position of selective realism is the kind of realism appropriate for scientific realism, rather than the uniform commitment to realism that is characteristic of, to take one example, Quine's criterion for ontological commitment.[37] Coupled with

[36] 'Observational' is used here in the liberal scientific sense, not in the philosophical sense.

[37] The 'realism' that results from the appearance of existential quantifiers in a theory is only a genuine form of realism when objectual quantification over domains of real objects is used. Quantification over domains of models, or other kinds of weaker interpretations of quantification are insufficient to establish genuine realism. Of course, using domains of real objects is, without clarification, blatantly circular. For an account of selective realism see Humphreys (2004), section 3.8.

Quine's resistance to second-order logic, the legacy of this criterion has distorted discussions of ontology. We can also be misled by too strong an emphasis on entity realism or statement realism, where entities are identified with objects, when an important part of the realism issue involves the existence of properties.

These are considerations relating to scientific realism and although empiricists and anti-realists have attempted to label specifically philosophical reasoning about such matters as illegitimate, such reasoning is to specifically ontological conclusions. Here is one further example that is indisputably metaphysical. Universal determinism can seem puzzling because it is committed to the view that the present state of the universe was fixed millions of years ago. The puzzlement can be made to go away, but there is a related puzzle that is less easy to resolve: How can all of the relevant information needed to fix future states of the universe be encoded into a single instantaneous state? This puzzle becomes more pressing when we consider the state at the first instant of the universe. How can everything needed to guide the development of the universe for evermore be right there in a single time slice which has no history and no law-like regularities to guide its subsequent development? This is a serious difficulty for any view that insists that there can be no causal properties without an associated law, and that regularities are a necessary condition of having a law. There are no non-trivial regularities at the start of the universe, hence no laws, and hence, on these views, no causes.

There is no difficulty for systems that can refer to rules. Computational systems are such that their dynamical development is driven by rules of the form: if the system is in such and such a state, then execute action thus and so. If we assume analogously that the laws of nature were there at the beginning, as part of the program for the universe, this answer immediately gives us an important realist conclusion, that the early stages of the universe contained laws that existed independently of specific states and of what we can know about them.

There are other approaches to addressing this puzzle. One says simply that there is no answer to our question, that there are just successive states of the universe, one after another, and parts of the first time slice can be attributed a causal role only if and when regularities appear. I note that this answer can be, and is, couched in realist terms rather than in the epistemic way that Humeans favor. If you feel that there are deep and probably unanswerable mysteries about the origin of the universe, then you should

feel that this first approach attributes those mysteries to every successive instant in the universe's evolution until a satisfactory regularity had been established. This is Hume's Problem dramatized, and it leaves the early and subsequent development of the universe a complete mystery. It is consistent but lacking in explanatory content.

A second approach denies the implicit assumption of the puzzle, which is that at the first instant there are no later instants yet in existence, and adopts the block universe model with its assumption that all states are eternally coexistent. This second approach has a curious feature, for if in such worlds laws supervene upon regularities and at least some different subsequent histories of the universe would give rise to different laws, and explanations of why later states of the universe have the features that they do depend upon both the initial state and the laws, then it is not true that the initial state of the universe determined (by itself) what was to come. Nor indeed does any initial segment of the universe, but only its entire subsequent development. This involves both the dependence of earlier states on later states and action at a temporal distance. A high price to pay and one that can be avoided by a third approach, which is to reject the idea that laws play a role in the development of the universe, and to assert that it is present property instances and their interactions that alone give rise to future developments. This is a metaphysical argument and one to which science itself has little if anything to contribute beyond providing constraints on the possible solutions.

8. Conclusions

The source of many of the problems associated with speculative ontology is that science, and physics in particular, long ago outran the conceptual abilities of most speculative ontologists. Contemporary science has revealed a much more subtle and interesting world than the often simple worlds of speculative ontologists, one example being the overuse of mereology by many contemporary metaphysicians.[38] The solution to these inadequacies is to pursue scientific ontology, an activity that is primarily interested in the

[38] There are scientifically sensitive advocates of more sophisticated forms of mereology. For one example, see Arntzenius and Hawthorne (2005) and for a balanced assessment of the mereological project see Varzi (2011).

contingent ontology of our world but one that can also provide a guide to unactualized possibilities. We can extract six morals from our discussion.

First, metaphysicians could, with a modest amount of effort, identify which contemporary metaphysical claims have been shown by science to be either contingently, rather than necessarily, true or simply false, thus preventing a considerable amount of wasted effort. It is true that advocates of speculative ontology can point to the pessimistic induction from the history of science as evidence that all scientific knowledge is eventually overthrown and argue that appeals to scientific authority are therefore too temporary and flimsy a basis for metaphysical truth. This would be an acceptable response if appropriate controls are placed on the methods used in metaphysics. A scientific ontologist must concede that fallibilism is the only acceptable epistemological position but our second conclusion, that different methods are appropriate for different domains of investigation, should be implemented. Thirdly, appeals to intuition are to be viewed with suspicion and it must at least be recognized that such appeals are both domain sensitive and differ in reliability between philosophers. Fourthly, conceptual analysis needs to rely less on ordinary concepts and draw more heavily on the extensive conceptual resources of science and, where appropriate, mathematics. Fifthly, philosophers collectively need to develop criteria for what counts as a legitimate philosophical idealization, criteria that are currently conspicuous by their absence. Finally, ontologists must recognize that our world is not scale-invariant and that as a result, a number of inferences that are taken to be a priori have a hidden inductive element.[39]

References

Arntzenius, F. and Hawthorne, J. (2005). Gunk and continuous variation. *The Monist*, 88: 441–65.

[39] I should like to thank the many philosophers with whom I have discussed the topics covered in this paper, especially the participants in a seminar at the University of Virginia in the spring of 2007, the audience at the University of Alabama conference on scientific naturalism and metaphysics in the fall of 2009, and lectures at the Ecole Normale Supérieure, Paris, the American Philosophical Association Pacific Division, and the University of Valparaiso, Chile. Comments on the penultimate draft from Anjan Chakravartty were particularly useful.

Bealer, G. (1998). Intuition and the autonomy of philosophy. In M. DePaul and W. Ramsey (eds.) *Rethinking Intuition*, pp. 201–39. Lanham, MD: Rowman and Littlefield.

——(1999). The a priori. In J. Greco and E. Sosa (eds.) *The Blackwell Guide to Epistemology*, pp. 243–70. Oxford: Blackwell.

Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of 'hidden variables', Parts I and II. *Physical Review*, 89: 166–93.

Carnap, R. (1956). Empiricism, semantics, and ontology. In *Meaning and Necessity*, pp. 205–21. Chicago, IL: University of Chicago Press.

Chalmers, D. (1997). *The Conscious Mind*. Oxford: Oxford University Press.

——and Jackson, F. (2001). Conceptual analysis and reductive explanation. *The Philosophical Review*, 110: 315–61.

Chittka, L. and Walker, J. (2006). Do bees like Van Gogh's sunflowers? *Optics & Laser Technology*, 38: 323–8.

De Millo, R., Lipton, R., and Perlis, A. (1979). Social processes and proofs of theorems and programs. *Communications of the ACM*, 22: 271–80.

DePaul, M. and Ramsey, W. (eds.) (1998). *Rethinking Intuition*. Lanham, MD: Rowman and Littlefield.

Dennett, D. (1991). Real patterns. *Journal of Philosophy*, 88: 27–51.

Feynman, R. (1967). *The Character of Physical Law*. Cambridge, MA: MIT Press.

Field, H. (1980). *Science Without Numbers*. Princeton, NJ: Princeton University Press.

Fine, A. (1984). The natural ontological attitude. In J. Leplin (ed.) *Scientific Realism*, pp. 83–107. Berkeley, CA: University of California Press.

Goldman, A. (2007). Philosophical intuitions: Their target, their source, and their epistemic status. *Grazer Philosophische Studien*, 74: 1–26.

Hawking, S. and Mlodinow, L. (2010). *The Grand Design*. London: Bantam Books.

Herzberg, F. (2010). The consistency of probabilistic regresses. A reply to Jeanne Peijnenburg and David Atkinson. *Studia Logica*, 94: 331–45.

Humphreys, P. (2004). *Extending Ourselves*. New York: Oxford University Press.

Jaffe, A. and Quinn, F. (1993). 'Theoretical Mathematics': Towards a cultural synthesis of mathematics and theoretical physics. *Bulletin (N.S.) of the American Mathematical Society*, 29: 1–13.

Jackson, F. (2000). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.

Johnson, M. (1997). Manifest kinds. *The Journal of Philosophy*, 94: 564–83.

Ladyman, J. and Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Lewis, D. (1986). *Philosophical Papers*, Volume II. Oxford: Oxford University Press.

Maudlin, T. (2007). *The Metaphysics within Physics*. Oxford: Oxford University Press.

Putnam, H. (1974). The meaning of 'meaning'. In K. Gunderson (ed.) *Language, Mind and Knowledge*, *Minnesota Studies in the Philosophy of Science*, vii: 131–93. Minneapolis: University of Minnesota Press.

Quine, W. (1969). Ontological relativity. In *Ontological Relativity and Other Essays*, pp. 26–68. New York: Columbia University Press.

Rosenhouse, J. (2009). *The Monty Hall Problem*. New York: Oxford University Press.

Stewart, D., Mackintosh, J., Playfair, J. et al. (1935). *Dissertation on the History of Metaphysical and Ethical, and of Mathematical and Physical Science*, pp. 580–602. Edinburgh: A & C Black.

Symons, J. (2008). Intuition and philosophical methodology. *Axiomathes*, 18: 67–89.

van Heijernoort, J. (ed.) (1967). *From Frege to Godel*. Cambridge, MA: Harvard University Press.

van Fraassen, B. (2004). *The Empirical Stance*. New Haven, CT: Yale University Press.

Varzi, A. (2011). Mereology. In E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*, Spring 2011 Edition [online]. <http://plato.stanford.edu/archives/spr2011/entries/mereology/> accessed 11 May 2012.

Weatherson, B. (2010). David Lewis. In E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*, Summer 2010 Edition [online]. <http://plato.stanford.edu/archives/sum2010/entries/david-lewis/> accessed 11 May 2012.

Williamson, T. (2005). Armchair philosophy, metaphysical modality, and counterfactual thinking. *Proceedings of the Aristotelian Society*, 105: 1–23.

4

Can Metaphysics Be Naturalized? And If So, How?

Andrew Melnyk

I began to study philosophy about 30 years ago.[1] It is clear to me, as it is to every philosopher who has lived through the intervening period, that the way in which philosophy is practiced today is very different from the way in which it was practiced then. The obvious outward sign of this difference in practice is the greatly increased probability that a philosophical journal article or book will discuss or cite the findings of some kind of empirical investigation, usually a science, but sometimes a branch of history. The difference itself is the (partial) so-called *naturalization* of many branches of philosophy.

Reflection on the contemporary practice of, say, philosophy of mind, philosophy of science, philosophy of language, moral philosophy, and even political philosophy suggests that the findings of empirical investigation play two main roles when philosophy is naturalized. First, they serve as evidence intended to confirm or disconfirm philosophical theses, theses that may themselves be quite traditional. For example, such findings have recently been used to cast doubt on the traditional claim that we have infallible knowledge of our own current experiences; and other findings to support an approximately Humean sentimentalism about moral judgments.[2] Second, such findings play the role of *object* of philosophical inquiry, in effect generating new philosophical questions. For example, the perplexing results of experiments performed on patients whose left and right cerebral hemispheres had been largely disconnected have generated much

[1] I.e., Anglophone philosophy, the only kind I know.

[2] See Schwitzgebel (2008) and Prinz (2006).

philosophical discussion of how to make sense of them.[3] More recently, the success of neuroscience and cell biology in discovering the mechanisms underlying various phenomena has prompted extensive efforts to understand what mechanistic explanation is.[4]

But when we turn to metaphysics as currently practiced by philosophers who think of themselves as metaphysicians (rather than as, say, philosophers of physics), we see no such signs of naturalization. Conforming to a long tradition, these metaphysicians do not cite empirical findings to confirm or disconfirm their contentions, and they do not address novel problems generated by such findings.[5]

It might be suggested that this is because there is no *reason* to naturalize metaphysics of this sort; it is a non-empirical inquiry that is managing quite well, thank you, as it is. The most promising way to develop this suggestion is to claim that metaphysics as currently practiced is, if not exactly a branch of mathematics or logic, then at least analogous to mathematics or logic, where these disciplines are understood traditionally, as a priori. But metaphysics as currently practiced is very unlike mathematics and logic. At any point in time, a remarkably broad consensus—amounting almost to unanimity—obtains among competent practitioners in mathematics and logic concerning the truth of a vast number of mathematical and logical claims. Moreover, the scope of this consensus has grown wider and wider over time, that is, competent practitioners at later times agree on more mathematical and logical truths than did competent practitioners at earlier times. In metaphysics, by contrast, we observe neither phenomenon; we instead observe persistent disagreement concerning what is true. This disagreement is strong prima facie evidence that contemporary metaphysicians do not have reliable methods for discovering metaphysical truths.

As I see it, then, the position is this: although metaphysics has not in fact been naturalized, it ought to be, since non-naturalized metaphysics has been pursued for a very long time without yielding results at all

[3] The pioneering discussion was by Thomas Nagel (1971).

[4] See, for example, Craver (2007).

[5] Admittedly, Jonathan Shaffer (2010) invokes quantum entanglement to support his metaphysical thesis of the priority of the whole (pp. 51–5). But in fact this invocation just emphasizes the gulf between his thesis and actual science. All that quantum entanglement shows is that the cosmos has properties that don't supervene on the intrinsic properties of, plus the interrelations among, its entangled parts. From this it doesn't follow that the cosmos is prior to its parts, *in Shaffer's sense of 'prior'*, unless there's such a thing as priority in his sense; and there is absolutely no scientific reason to believe that there is.

comparable with those achieved by mathematics and logic. But *can* metaphysics be naturalized? And if it can, how exactly can the results of empirical investigation be made relevant to metaphysics—as evidence, as a source of new problems, or in other ways? How many traditional metaphysical problems will it still be reasonable to investigate? And if the answer is 'Not many,' then what sort of problems will take their place? These are the sorts of questions that I wish to explore in this paper.

I do not approach these questions, however, with the assumption that metaphysics is bound to turn out to be a viable branch of inquiry, and hence that the only live question is how it works. On the contrary, I think there is a real possibility that the activity that we call 'metaphysics' should turn out not to constitute a viable form of inquiry at all, either empirical or non-empirical. I am therefore prepared to find that the right answer to the question, 'Can metaphysics be naturalized?' is 'No, it can't.'

My procedure in what follows will be slightly unorthodox. I will allow my answers to the questions I have raised to emerge from close dialogue with the first chapter (co-authored by Don Ross, James Ladyman, and David Spurrett; henceforth 'RLS'[6]) of Ladyman and Ross's remarkable book, *Every Thing Must Go: Metaphysics Naturalized*.[7] I single out this chapter for such full examination because it is far and away the richest account to date (i) of why mainstream analytic metaphysics is objectionably non-naturalistic and (ii) of how metaphysics might be naturalized. I will be quoting from it liberally. I find RLS's critique of mainstream analytic metaphysics very powerful, but I have significant reservations about their positive conception of naturalized metaphysics, as I shall explain.

I

RLS begin their chapter as follows:

> The aim of this book is to defend a radically naturalistic metaphysics. By this we mean a metaphysics that is motivated exclusively by attempts to unify hypotheses and theories that are taken seriously by contemporary science. (1)

[6] Since 'RLS' simply abbreviates the names of the three co-authors, I treat it as grammatically plural.

[7] All page references that follow are to this book, unless the contrary is clearly indicated.

But their speaking of '*a* radically naturalistic metaphysics' (italics added) does not indicate tolerance for other kinds of metaphysics, for they immediately add this:

> For reasons to be explained, we take the view that no alternative kind of metaphysics can be regarded as a legitimate part of our collective attempt to model the structure of objective reality. (1)

Even these few remarks make it clear that not much of what contemporary analytic philosophers do under the heading of 'metaphysics' counts as legitimate by RLS's lights.

In my next section, I shall look at the details of RLS's conception of naturalized metaphysics. In this section, I shall ask three general questions about it that don't require knowing the details.

My *first* question about RLS's conception of naturalized metaphysics is why they think that *unification* is the touchstone of legitimate metaphysics. On the strength of their pp. 27–8, I think their answer can be paraphrased as follows:

> The goal of 'a relatively unified picture of the world' is pursued by actual scientists—and rightly so. But unifying whole branches of science, by figuring out the 'reciprocal explanatory relationships' between them, 'is not a task assigned to any particular science'. Doing so, therefore, is a truth-oriented task for metaphysics that is not crowded out, so to speak, by any single science.

I am all in favor of seeking 'a relatively unified picture of the world';[8] but I have two reservations about this argument.

First, I don't see why this argument should provide any reason to *restrict* naturalized metaphysics to attempts at unification. After all, the goal of discovering true answers to questions we find important is *also* a goal pursued by actual scientists; and some of these questions are also not addressed by any single branch of science. Examples are the questions whether anything ever causes anything, whether the world is fundamentally impersonal, whether anything has intrinsic properties, or indeed whether there are any fundamental individuals.[9] Thus there is, I suggest, as good a

[8] See the 'problem of the many sciences' in Andrew Melnyk (2003).

[9] Readers familiar with *Every Thing Must Go* will know that its Ch. 3 argues in detail that what contemporary physics, properly viewed, is really telling us is that there are no fundamental individuals. While this chapter is in fact integrated into the book's overall project of giving a unified account of the world as a whole, it would still, I think, be a legitimate contribution to an acceptably naturalized metaphysics even if it were not (and even if it

rationale for allowing naturalized metaphysics to seek answers to these questions as there is for allowing it to seek global unification.

Second, although RLS are right that no single branch of science is tasked with generating an account of how all the branches of science fit together, there are branches of science, for example, physical chemistry and molecular biology, that give accounts of the relations between members of particular pairs of branches of science—and that do so without philosophical assistance. The question then arises of what would be wrong with simply conjoining all these accounts (including future ones) and letting the result *be* one's account of how the sciences and their respective domains are to be unified. Why would this merely conjunctive unifying account of the world not crowd out naturalized metaphysics as RLS envisage it, leaving it with nothing to do, even if no single science crowds it out?

A plausible answer to this question, I suggest, is that the accounts of the relations between branches of science provided by, for example, physical chemistry and molecular biology are *deficient* in some way that philosophers are in a position to remedy. The best candidate for such deficiency, however, is not the obvious one, that is, that these accounts are *false*, but rather that they are *imprecise*. I shall not try to define imprecision, but the kind I have in mind is exemplified by the pervasive claims in cognitive neuroscience that such-and-such a neural condition is the 'neural basis' of such-and-such a psychological state or capacity. When such claims are made, no account of *being the basis of* is ever offered at the time, and no consensus account can be assumed, since none exists.

The thought that some of the products of science might be deficient in some way that philosophers could remedy leads nicely into my *second* question about RLS's conception of naturalized metaphysics. According to this conception, as we have seen, the only legitimate metaphysics is one 'that is motivated exclusively by attempts to unify hypotheses and theories that are taken seriously by contemporary *science*' (1; my italics). But why do RLS privilege science in this way? After all, one could agree that legitimate metaphysics must be some kind of unification project, but deny that the claims about the world that it seeks to unify be drawn only from science.

couldn't be; this would be the case if there simply was no unified account of the world as a whole).

RLS offer the following argument for the unique status of science; they repeat it at p. 30:

> Since science just *is* our set of institutional error filters for the job of discovering the objective character of the world—that and no more but also that *and no less*—science respects no domain restrictions and will admit no epistemological rivals (such as natural theology or purely speculative metaphysics). With respect to anything that is a putative fact about the world, scientific institutional processes are absolutely and exclusively authoritative. (28)

Their rationale for the premise of this argument is that scientific methods—of discovery and confirmation—are at bottom no different in kind from the methods of discovery and confirmation used in everyday life, for example, in police work, courts of law, or auto mechanics; scientific methods are simply the most refined, best developed, and most effective of these everyday methods.

Now if scientific methods are indeed the best refinements to date of certain everyday methods for acquiring knowledge, which I accept, it certainly follows that, in any *conflict* between the deliverances of scientific methods and the deliverances of those everyday methods, we should always prefer the former. But does it also follow that scientific methods are 'exclusively authoritative,' that is, that no other methods can bear on, or even settle, a factual claim? Apparently not. For all that has been said so far, scientific methods may fail to include refinements or developments of *all* everyday methods of inquiry that have cognitive value. (Why expect that they would include them all? What mechanism would ensure this outcome?) And defenders of, say, a somewhat reliable faculty of introspection, or indeed of a *sensus divinitatis*, will obviously claim that in fact they do so fail. It also doesn't follow, from the claim that scientific methods are the best refinements to date of certain everyday methods for acquiring knowledge, that everyday claims about the world shouldn't be taken seriously unless they have been vindicated by the procedures of institutional science. For, in addition to the last point, everyday methods of inquiry, without being the *best* methods we have, may still be at least somewhat reliable.

I am inclined, then, to reject RLS's conclusion—that 'With respect to anything that is a putative fact about the world, scientific institutional processes are absolutely and exclusively authoritative'—as unjustifiably strong. I also doubt that any such conclusion can be established by means of the sort of global argument that RLS offer. Instead, one must argue

piecemeal, by evaluating concrete considerations for and against whatever particular methods of inquiry might be proposed as supplements to those of science. For example, the deliverances of intuition concerning certain factual claims could be evaluated for coherence with one another, both at a time and over time, both within subjects and between subjects; and for external coherence, that is, coherence with the deliverances of other sources of evidence already accepted as legitimate. Such deliverances can also be evaluated by seeking a theoretical account of their origins and reliability. In practice, RLS do argue in this piecemeal way, as when they point out that intuitions are often influenced by variable cultural factors or superseded scientific theories, and that common-sense claims have often turned out to be wrong or unproductive (10–12).

My *third* question about RLS's conception of naturalized metaphysics is closely related to the second. What scope does it allow for metaphysics, and indeed for philosophy more broadly, to *correct* our best current science? The question matters rhetorically as well as substantively. It matters substantively because, as I've hinted already, a *prima facie* tension exists between the naturalist attitude of deferring to science on all factual questions and the hope that metaphysics can nonetheless contribute to our knowledge of what the world is like. It matters rhetorically because naturalists are in constant danger of appearing to be science sycophants, and they would reduce this threat if they could point to ways in which, in principle, metaphysicians could correct science.

Now when RLS write that 'With respect to anything that is a putative fact about the world, scientific institutional processes are absolutely and exclusively authoritative,' they certainly appear committed to allowing no scope at all for metaphysics to correct science. But the appearance is misleading. Suppose that the methods of properly naturalized metaphysics *are* (some of) those of science; this involves supposing that properly naturalized metaphysics has (or could be made to have) its own 'institutional filters on errors' (28), comparable to those of today's science. Then, in principle, no obstacle prevents such metaphysics from correcting our best current science, and all sorts of possibilities are opened up. For example, properly naturalized metaphysics might be able to show that parts of our best current science are imprecise (as I suggested above), or confused, or needlessly agnostic; that these defects can be corrected in such-and-such a way; and that, once they have been corrected, we will

therefore have *added* to our current best science.[10] Properly naturalized metaphysics might also be able to show that parts of our best current science are *unfounded*, because they are supported by faulty modes of reasoning; in this case, properly naturalized metaphysics would presumably need to draw upon properly naturalized epistemology. More generally, practitioners of properly naturalized metaphysics might be able to advance our best current science indirectly, merely by helping scientists to think things through, without their contribution amounting to a precisely identifiable, localized addition to anything that we take ourselves to know.

II

I turn now to RLS's elaboration of their conception of naturalized metaphysics. In a section devoted to formulating 'some proper principle which distinguishes what we regard as useful from useless metaphysics' (27), they endorse a 'non-positivist version of verificationism' (29). They state the first of the two elements that make up this 'verificationism' as follows:

> no hypothesis that the approximately consensual current scientific picture declares to be beyond our capacity to investigate should be taken seriously. (29)

RLS make it clear in their glosses that 'capacity' should be read as 'capacity in principle'; and that not to take a hypothesis seriously is to treat the hypothesis as one whose investigation is not worthwhile if one's (sole) goal is to advance 'objective inquiry' (30). Note, too, that the sentence I have quoted is an exhortation, and not a hypothesis, so that no problem can arise from its applying to itself.

But why call this exhortation an element of *verificationism*? The answer is that classical verificationism sought a way to identify claims that are cognitively meaningless, claims that can for that reason safely be ignored, that is, left uninvestigated. And RLS, too, seek a way to identify hypotheses that can safely be ignored—not because they are literally meaningless, but because they are beyond our cognitive powers to investigate. A second query: what does RLS's exhortation have to do with metaphysics? The answer is everything, since 'no hypothesis' is clearly meant to cover all

[10] Indeed, this is just what will happen, I suggest, if the arguments of Ch. 3 of *Every Thing Must Go* (that according to current physics there are no fundamental individuals) are successful.

metaphysical theses. The exhortation is therefore closely linked to a certain familiar usage of the word 'metaphysical' (chiefly among scientists) in which to call a claim metaphysical is just to say that its truth can't be known, one way or the other, even in principle. Just how many traditional metaphysical claims should not be taken seriously, according to RLS, is a question they don't take up at this point in their discussion; but their earlier section, 'Neo-Scholastic Metaphysics' (7–27), can with the benefit of hindsight be read as having argued precisely that a good many such claims are indeed 'beyond our capacity to investigate', given the 'approximately consensual current scientific picture' insofar as it characterizes human cognition. Thus, as I noted above, RLS observe that intuitions are often influenced by variable cultural factors or superseded scientific theories, and that common-sense claims have often turned out to be wrong or unproductive (10–12). Hence, to the extent (i) that certain metaphysical theses could only be supported by appeal to intuitions and to what seems commonsensical and (ii) that RLS's observations discredit support of these kinds, those theses should no longer be taken seriously, according to the first element of RLS's verificationism.

For myself, I am not yet ready to endorse (ii). The evidence that RLS cite against appeals to intuitions and common sense is too general. A great diversity of phenomena has been subsumed under the heading of 'intuitions and common sense,' and it is open to defenders of the appeals to intuitions and common sense that are made in contemporary metaphysics to insist that the sort of appeals to intuitions and common sense that *they* make are legitimate, even though RLS are right to say that many other appeals that fall into the same very broad category are worthless. In my view, an adequate case against the appeals to intuitions and common sense that are made in contemporary metaphysics must be based on the results of further research into metaphysical intuitions in particular (e.g., those invoked in the dispute between endurantism and perdurantism). Such research would seek, first, to discover whether metaphysical intuitions are consistent across, say, cultures, genders, and variations in intelligence and education; in light of the results, it would seek, second, to confirm hypotheses about where these intuitions come from and what factors influence them; it would seek, finally, to draw conclusions as to the reliability or otherwise of these intuitions.

Even when the 'approximately consensual current scientific picture' doesn't characterize human cognition, it can still reveal a hypothesis to be 'beyond our capacity to investigate'; it can do so by revealing certain

questions to have false presuppositions. Thus RLS see much contemporary analytic metaphysics as assuming the correctness of what they call the 'containment metaphor' (3). 'On this doctrine,' they write,

> the world is a kind of container bearing objects that change location and properties over time. These objects cause things to happen by interacting directly with one another. Prototypically, they move each other about by banging into one another. At least as important to the general picture, they themselves are containers in turn, and their properties and causal dispositions are to be explained by the properties and dispositions of the objects they contain (and which are often taken to comprise them entirely). (3)

According to the balance of *Every Thing Must Go*, however, this doctrine is entirely mistaken; most fundamentally, of course, there are no objects, that is, no substances in the philosophical sense—every thing must go! But if this is correct, then, I take it, many traditional metaphysical questions simply lapse, such as questions about what substances are, how they continue to exist through time, whether some of their properties are essential, and perhaps what properties are (if substance and property are inter-defined). For RLS, in that case, naturalized metaphysics *cannot* mainly consist in addressing traditional metaphysical questions, albeit in a non-traditional way; for them, the change that metaphysics must undergo if naturalized is more than a change in method.

Let me end this section with a word about the rationale for the first element of RLS's 'non-positivist version of verificationism.' Why, exactly, should 'no hypothesis that the approximately consensual current scientific picture declares to be beyond our capacity to investigate . . . be taken seriously' if one's (sole) goal is to advance 'objective inquiry' (29–30)? Though left unstated, the answer, I presume, is that—to a first approximation—one should not attempt to pursue a goal that one knows one cannot achieve. A more sophisticated answer would need to take into account the low but surely non-zero probability that one *can* achieve the goal in question, *pace* the 'approximately consensual current scientific picture,' which is, after all, fallible. Perhaps, then, despite the very low probability of success, if some people want very badly to know whether, for example, endurantism or perdurantism (or neither) is true, they *should* continue to investigate them as best they can. But RLS can happily concede this, for their main points stand: (i) it is sometimes irrational for most people to continue certain investigations because the game is not worth the candle, given their utilities and the

low probability of discovering, or even approaching or approximating, the truth, and (ii) contemporary metaphysics may well be an investigation of this kind.[11]

III

The second of the two elements that make up RLS's 'non-positivist version of verificationism', once developed and refined, is what they call their 'principle of naturalistic closure' or 'PNC' (27) and formulate canonically as follows (37–8):

> Any new metaphysical claim that is to be taken seriously at time *t* should be motivated by, and only by, the service it would perform, if true, in showing how two or more specific scientific hypotheses, at least one of which is drawn from fundamental physics, jointly explain more than the sum of what is explained by the two hypotheses taken separately, where this is interpreted by reference to the following terminological stipulations.

These stipulations explain the intended senses of 'scientific hypothesis' and 'specific scientific hypothesis.'

RLS clearly intend the PNC as an elaboration of the general idea with which they begin their chapter, that 'a radically naturalistic metaphysics . . . [is] a metaphysics that is motivated exclusively by attempts to unify hypotheses and theories that are taken seriously by contemporary science' (1). So my earlier comments on the general idea apply to the PNC as well. But the PNC differs from the general idea in two striking ways. First, it includes mention of a particular branch of science, fundamental physics. Second, it appears no longer to mention unification. Let me begin by discussing the fate of unification in the PNC; on my interpretation, it's still there, but it's harder to see.

What exactly do RLS mean when they say that 'a metaphysical claim' could perform a service 'in showing how two or more specific scientific hypotheses . . . jointly explain more than the sum of what is explained by the two hypotheses taken separately'? If the quoted formulation of the PNC were read out of context, its talk of jointly explaining 'more' would most naturally be construed as talk of jointly explaining a greater *number* of

[11] In my view, certain exegetical questions that have been discussed by historians of philosophy for literally (two) thousands of years may also fall into this category.

explananda. But RLS see themselves as borrowing from Philip Kitcher's well-known account of explanatory unification (30–2). So their talk of jointly explaining 'more' should be construed instead as referring to a gain in explanatory power—something that does not necessarily require a greater number of explananda. Thus RLS paraphrase Kitcher's account, with endorsement, as follows (roughly, argument patterns are multiply-applicable explanatory schemata that consist of schematic sentences):

> We have a unified worldview to the extent that [i] we use a smaller rather than a larger number of argument patterns in science, and to the extent that [ii] what get used as schematic sentences in these argument patterns are themselves derived from other non-ad hoc argument patterns. (31; interpolated numerals mine)

Given this account, if two hypotheses 'jointly explain more than the sum of what is explained by the two hypotheses taken separately', then, I suggest, this must be because the two hypotheses 'jointly' constitute 'a smaller rather than a larger number of argument patterns'; and presumably the two hypotheses 'jointly' constitute 'a smaller rather than a larger number of argument patterns' because the argument patterns of one of the hypotheses 'are themselves derived from other' argument patterns that belong to the other hypothesis. I further suggest that 'a metaphysical claim' could perform a service in showing *how* two hypotheses 'jointly explain more than the sum of what is explained by the two hypotheses taken separately' by showing *how* the argument patterns of one of the hypotheses 'are themselves derived from other' argument patterns that belong to the other hypothesis; and showing this would presumably require affirming some substantive connection between the two hypotheses and/or their respective domains.

As noted above, the PNC requires that one of the two (or more) hypotheses be 'drawn from fundamental physics.' At first sight, this requirement seems too strong: one might have thought that unifying 'two or more specific scientific hypotheses,' even if *neither* of them was 'drawn from fundamental physics,' could still be one step on the path to a later unification that *did* involve a hypothesis 'drawn from fundamental physics.' But RLS insist upon the requirement, writing as follows:

> a hypothesis that unified specific hypotheses from sciences other than fundamental physics, but unified them with no specific hypotheses from fundamental physics, would not be a metaphysical hypothesis. It would instead be a hypothesis of a special science of wider scope than those it partially unified. (37)

The reason why it wouldn't be 'a metaphysical hypothesis' is that—if I have it right—naturalized metaphysics should 'share . . . the maximum scope of fundamental physics' (37); and it should 'share . . . the maximum scope of fundamental physics' because of an important principle that RLS call the 'Primacy of Physics Constraint' (37).

I shall make just one comment on this argument before I turn to the Primacy of Physics Constraint (or PPC). Let me grant that naturalized metaphysics should 'share . . . the maximum scope of fundamental physics.'[12] I still don't see why the unification of a specific scientific hypothesis with a specific hypothesis from fundamental physics *would* be metaphysical, whereas the unification of two specific hypotheses *not* drawn from fundamental physics would *not* be. *Neither* unification *actually* achieves maximum scope; and *both* seem equally good candidates to be *precursors* to the achievement of maximum scope, namely, the unification of all science.

IV

Here is how RLS formulate the PPC:

> Special scientific hypotheses that conflict with fundamental physics, or such consensus as there is in fundamental physics, should be *rejected* for that reason alone. Fundamental physical hypotheses are not symmetrically hostage to the conclusions of the special sciences. (44; my italics)

Why do RLS accord fundamental physics this special epistemic status? Because, they say, the history of science from the nineteenth century onwards 'has been widely taken to support two complementary arguments for the primacy of physics' (42–3). According to the first argument,

> in the history of science a succession of specific hypotheses to the effect that irreducibly non-physical entities and processes fix the chances of physical outcomes have failed. (43)

But so what? I agree that the history of science provides inductive evidence that there are no irreducibly non-physical influences on the chances of physical outcomes. But it doesn't follow that the special sciences should always yield to physics in the event of a conflict. If, in defiance of the PPC,

[12] I grant this only *arguendo*, since I don't see why naturalized metaphysics should 'share . . . the maximum scope of fundamental physics,' even if the PPC (explained below) is endorsed.

we treat a conflict between a special-scientific hypothesis and the consensus view in fundamental physics as evidence that the *physics* is in error, we're not thereby committed to irreducibly non-physical influences on the chances of physical outcomes. Rather than explaining the conflict by supposing that physicists have overlooked certain irreducibly non-physical influences on the chances of physical outcomes, we could instead explain it by supposing that they have misunderstood some of the *physical* influences on the chances of physical outcomes.

According to the second argument for the primacy of physics,

> Over the history of science a succession of processes in living systems, and in the parts of some living systems dedicated to cognition, have come to be largely or entirely understood in physical terms. (43)

But this argument falls to the same objection as the first argument did. Let's agree that the history of science provides inductive evidence that *all* processes in living systems can in principle be entirely understood in physical terms. If we treat a conflict between a special-scientific hypothesis and the consensus view in fundamental physics as evidence that the physics rather than the hypothesis is in error, we needn't be assuming that some processes in living systems *can't* be entirely understood in physical terms. The conflict could have arisen because physicists have misunderstood the *physical* factors in terms of which processes in living systems can in principle be understood.

A page or two later, RLS offer a different (but consistent) rationale for the PPC. They claim that the PPC

> is a regulative principle in current science, and it should be respected by naturalistic metaphysicians. The first, descriptive, claim is reason for the second, normative, one. (44)

How exactly is the reasoning here meant to go? Perhaps as follows. As we saw above, RLS think that 'scientific institutional processes are absolutely and exclusively authoritative' (28). If this view is correct, and if respect for the PPC is one of the *components* of 'scientific institutional processes', then every reasonable person should respect the PPC. But is it true that respect for the PPC is one of the components of 'scientific institutional processes'? The evidence that RLS cite is inconclusive, since it doesn't discriminate between the hypothesis that scientists hold themselves to the PPC and the rival hypothesis that they hold themselves to a logically weaker relative of the PPC, the weaker relative claiming that conflict between a

special-scientific hypothesis and such consensus as exists in fundamental physics is *defeasible evidence* against, albeit perhaps *strong* defeasible evidence against, the special-scientific hypothesis. What makes this relative weaker, of course, is that it allows for the possibility that the evidence *for* the special-scientific hypothesis might outweigh the evidence *against* it that is constituted by its conflict with fundamental physics—in which case, contrary to the PPC, the special-scientific hypothesis should *not* be rejected. Scientists who clearly wish their special-scientific hypotheses not to conflict with physics might be implicitly endorsing not the PPC but its weaker relative.

Let me conclude this section with a comment about the practical employment of the PPC. In order for the PPC to recommend the rejection of a special-scientific hypothesis, a special-scientific hypothesis must be capable of *conflicting* with some claim of fundamental physics. Now if physicalism is assumed, then such a conflict is certainly possible, because, if physicalism is true, then a complete description, in the language of physics, of the physical way things are, plus all the true a posteriori identity claims, entails all the positive truths expressible in the languages of the special sciences, which truths have the potential to be contradicted by claims generated by the actual special sciences. But RLS deny the thesis of physicalism in its commonest formulations, and refuse to commit themselves to physicalism in any formulation (38–41). So, for them, how can it happen that a special-scientific hypothesis conflicts with some claim of fundamental physics?

I am not suggesting that RLS's position (that physical hypotheses might conflict with special-scientific hypotheses, even though physicalism is false) is incoherent. Indeed, it is plainly coherent. For if a complete description, in the language of physics, of the physical way things are, plus all the true a posteriori identity claims, entailed some *but not all* of the positive truths expressible in the languages of the special sciences, then physicalism would be false, but the entailed truths would still have the potential to conflict with claims generated by the actual special sciences. However, this possibility doesn't help much, because it still assumes that a correct physical description of the way things are entails *some* of the positive truths expressible in the languages of the special sciences, and even this weaker assumption, I think, is one that RLS would wish to withhold commitment to. So they still owe an explanation of how a special-scientific hypothesis could conflict with some claim of fundamental physics, and that is all I wished to note.

V

I have argued, amongst other things, that the core of RLS's position—the PNC—is too restrictive, and that their deference to institutional science is exaggerated. So I don't think that metaphysics can be naturalized in exactly the way that they propose. But I am cautiously optimistic that metaphysics can be naturalized. First, the optimism. In order for metaphysics to be naturalized, there have to be some outstanding questions that (i) we would like to answer but that (ii) apparently don't fall within the province of the sciences (as traditionally understood). There look to be such questions, including (but not limited to) the question of how to unify the sciences.

Second, however, the caution. In order for metaphysics to be naturalized, these questions have to be ones that creatures like us are capable in principle of answering. But if creatures like us really are capable of answering these questions, then, since a priori methods have a good track record only with regard to the mathematical and logical domains, our ability must derive from the application to the questions of empirical methods. But how far empirical methods can be applied to these questions is not clear to me. If a metaphysical question can be put into the 'What is__?' form (e.g., 'What is causation?'), then in principle it can be answered by assembling empirical evidence for the relevant a posteriori identity claim.[13] And the same approach can be used to address the question of how to unify the sciences, since, at least on my view, unification is achieved by discovering cross-scientific a posteriori identity claims.[14] So far, so good. But what if our question is, say, whether there are any fundamental substances? On the face of it, the only possible approach to such a question requires scrutinizing our best current physical theories and working from there—which is exactly what Ladyman and Ross do in Chapter 3 of *Every Thing Must Go*. Here, however, I have a concern, and it is one that RLS explicitly note: 'science, usually and perhaps always, underdetermines the metaphysical answers we are seeking' (9). Perhaps, then, one can get from the science only as much metaphysics as one puts in. Perhaps. But it's best not to meet one's troubles halfway. If

[13] For defence, see the third section of my 'Conceptual and Linguistic Analysis: A Two-Step Program' (2008).

[14] This is controversial, of course; RLS would deny it.

metaphysics cannot be naturalized, let us discover the fact naturalistically—by trying our best to do it and failing nonetheless.[15]

References

Craver, C. (2007). *Explaining the Brain: What the Science of the Mind-Brain Could Be*. Oxford: Oxford University Press.

Ladyman, J. and Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Melnyk, A. (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. New York: Cambridge University Press.

—— (2008). Conceptual and linguistic analysis: A two-step program. *Noûs*, 42, 2: 267–91.

Nagel, T. (1971). Brain bisection and the unity of consciousness. *Synthese*, 22: 396–413.

Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, 9: 29–43.

Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review*, 117: 245–73.

Shaffer, J. (2010). Monism: The priority of the whole. *Philosophical Review*, 119: 31–76.

[15] An ancestor of this paper was presented at a session of the Society for the Metaphysics of Science held at the annual meeting of the Pacific Division APA in San Diego in April 2011. I owe thanks to those present for their useful comments.

5

Kinds of Things—Towards a Bestiary of the Manifest Image

Daniel C. Dennett

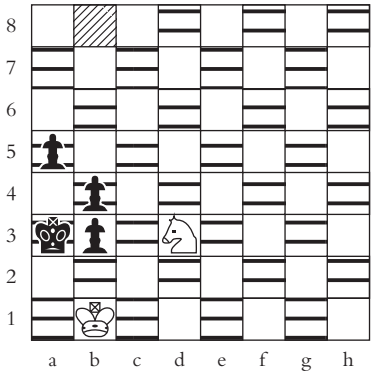

Figure 5.1 White to play (hint: checkmate in two moves).

Consider this chess puzzle. White to checkmate in two. It appeared recently in the *Boston Globe*, and what startled me about it was that I had thought it had been proven that you can't checkmate with a lone knight (and a king, of course). I was wrong; as David Misialowski (personal correspondence) has pointed out to me, what had been proven is just that you cannot checkmate your opponent's king when only his king and your king and knight are on the board. The fact that the proposition *you can never checkmate with a lone knight and king* is not a truth of chess is a *higher-order* truth of chess. Now let's consider chmess (I made up the term); it's the game that you get by allowing the king to move two spaces, not one, in any direction.

I made it up and I don't know if anybody has ever played it. I have no idea whether it's worth playing. Probably it isn't, but it doesn't matter; it's not even worth our attention long enough to figure out. But a moment's reflection reveals that there are exactly as many higher-order a priori truths of chmess as there are of chess, namely, an infinity of them. And no doubt they would be roughly as difficult to discover and to prove as the higher-order truths of chess. There are people who make a living working out the truths of chess and certainly it's been a big avocation for many other people. But I doubt if anybody yet has spent more than five minutes trying to work out the a priori truths, and the higher-order truths, of chmess.

I use this example (complete with my mistake, now corrected by Misialowski) in a paper, 'Higher-order truths about Chmess' (Dennett, 2006), my own salvo against analytic metaphysics and other dubious battles in philosophy. It is, more broadly, my advice to graduate students in philosophy on how to avoid the trap of getting stuck in a career devoted to an artifactual 'problematic.' As one of my heroes, Donald Hebb once said, 'If it isn't worth doing, it isn't worth doing well.' I think we have to admit that there are quite a few philosophers who do very well what they do, executing their self-assigned tasks with cleverness, technical proficiency, and rigor, but engaged in an enterprise that's just not worth doing.

So I applaud the first chapter of Ladyman and Ross (2007), which articulates the grounds for my conviction much better than I could ever have done. They comment briefly on my discussion (in Dennett, 2005) of Patrick Hayes, the artificial intelligence researcher, who set out on a project to axiomatize the naïve (or folk) physics of liquids. I'm going to expand a bit on this point, to clarify the message we both derive from this ingenious enterprise. The idea was that a robot who interacted with everyday folk would need to understand folk physics—the manifest image of liquids, solid objects, and so on—so, in classic GOFAI[1] style Hayes set out provide such a robot with the propositions it would need to use as its core beliefs. It proved to be more challenging than he had anticipated, and he wrote an interesting paper about the project, 'The Naïve Physics Manifesto' (Hayes, 1978). In the naïve physics of liquids, everything that strikes naïve folks as counter-intuitive is, of course, ruled out: siphons are 'impossible' and so are pipettes, but you can mop up liquid

[1] Good Old-Fashioned Artificial Intelligence.

with a fluffy towel, and pull water out of a well with a suction pump. A robot equipped with such a store of 'knowledge' would be as surprised by a siphon as most of us were when first we saw one in action. Hayes' project was what I would call *sophisticated* naïve physics, because he was under no illusions; he knew the theory he was trying to axiomatize was false, however useful in daily life. This was an exercise in what might be called axiomatic anthropology: you treat what the folk say—and agree about—as your axioms or theorems, and try to render the data set consistent. And, of course, he didn't bother rounding up any actual informants; he figured that he knew the naïve physics of liquids as well as any normal person did, so he used himself as his sole informant: axiomatic *auto*-anthropology.[2]

Now compare Hayes' project with the projects of analytic metaphysics, which often strike me as *naïve* naïve auto-anthropology since the participants in this research seem to be convinced that their program actually gets at something true, not just believed-true by a particular subclass of human beings (Anglophone philosophers of the analytic metaphysics persuasion). Otherwise, the programs seem identical: you gather your shared intuitions (testing and provoking them by engaging in mutual intuition-pumping) and then try to massage the resulting data set into a consistent 'theory,' based on 'received' principles that count, ideally, as axioms. I've asked a number of analytic metaphysicians whether they can distinguish their enterprise from naïve auto-anthropology of their clan, and have not yet received any compelling answers.

The alternative is *sophisticated* naïve anthropology (both auto- and hetero-)—the anthropology that reserves judgment about whether any of the theorems adduced by the exercise deserves to be trusted—and this is a feasible and frequently valuable project. In this essay I am going to propose that this is the enterprise to which analytic metaphysicians should turn, since it requires rather minimal adjustments to their methods and only one major revision of their *raison d'être*: they must roll back their pretensions and acknowledge that their research is best seen as a

[2] Hayes' work inspired others. See the anthology edited by Daniel Bobrow, *Qualitative Reasoning about Physical Systems* (1985). Authors of one of the essays therein remark: 'Naïve physics is in itself an ambiguous term. Is it just bad physics? Is it psychology? Artificial intelligence? Physics?' (de Kleer and Brown, p. 13) The answer, I submit, is that naïve physics is the attempt to make the physics part of our manifest image rigorous enough to support automated deductive reasoning.

preparatory reconnaissance of the terrain of the manifest image, conducted between brackets, as it were, like the Husserlian *epoché*: let's pretend for the nonce that the natives are right, and see what falls out. Since at least a large part of philosophy's task, in my vision of the discipline, consists in negotiating the traffic back and forth between the manifest and scientific images, it is a good idea for philosophers to analyze what they are up against in the way of folk assumptions before launching into their theory-building and theory-criticizing. Philosophical work on the perennially hot-button topic of free will, for instance, certainly must be guided by an appreciation of what non-philosophers think free will is or might be—and why it matters so much to them.

One of the hallmarks of sophisticated naïve anthropology is its openness to making counterintuitive discoveries. As long as you're doing naïve anthropology, counterintuitiveness (to the natives) counts against your reconstruction; when you shift gears and begin asking which aspects of the naïve 'theory' are true, counterintuitiveness loses its dispositive force and even becomes, on occasion, a sign of significant progress. In science in general counterintuitive results are prized, after all.

One of the weaknesses of auto-anthropology is that one's own intuitions are apt to be distorted by one's theoretical predilections. Linguists have known for a long time that they get so wrapped up in their theories that they are no longer reliable sources of linguistic intuition. Their raw, untutored intuitions have been sullied by too much theory, so they recognize that they must go out and ask non-linguists for their linguistic intuitions. Philosophers have recently begun to appreciate this point, in the new enthusiasm for so-called experimental philosophy. It is early days still, and some of the pioneer efforts are unimpressive, but at least philosophers are getting used to the idea that they can no longer just declare various propositions to be obviously true on the sole grounds that they seem smashingly obvious *to them*. (In a similar vein Hayes *might* have surprised himself about the chief tenets of folk physics if he had gone to the trouble of interviewing a random sample of folk instead of just treating himself as exemplary.)

So there is a project, a particular sort of sophisticated naïve anthropology, that philosophers should seriously consider as a propadeutic inquiry before launching into their theories of knowledge, justice, beauty, truth, goodness, time, causation, and so on, to make sure that they actually aim their analyses and arguments at targets that are relevant to the rest of the

world, both lay concerns and scientific concerns. What they would get from such a systematic inquiry is something like a *catalogue raisonné* of the unreformed conceptual terrain that sets the problems for the theorist, the metaphysics of the manifest image, if you like. This is where we as philosophers have to start in our attempts to negotiate back and forth between the latest innovations in the scientific image, and it wouldn't hurt to have a careful map of this folk terrain instead of just eyeballing it. (This is the other half, one might say, of the reform that turned philosophy of science from an armchair fantasy field into a serious partnership with actual science when philosophers of science decided that they really had to know a lot of current science from the inside.) Once one thinks about our philosophical tasks with this image in mind, one can see that a great deal of the informal trudging around, backing and filling, counter-example-mongering and intuition-busting that fills the pages of philosophy journals is—at best—an attempt to organize a mutually acceptable consensus about this territory. Maybe a Mercator of the Manifest Image will emerge from experimental philosophy.

But that won't be me. I am instead going to fall back on the good old-fashioned informal methods of ordinary language philosophy—the 'discipline' I was 'trained' to do—and draw attention to some of the unanswered questions that one might want to take seriously in sketching the ontology of the manifest image. The only innovation in my method is that I encourage us all to be *sophisticated* naïve (auto-) anthropologists, perhaps *savoring* our intuitions about cases, but not *endorsing* them—yet.

The main goal in the rest of this paper is simply to draw attention to the riotous assembly of candidate *things* that we find in the manifest image. Wilfrid Sellars (1963) 'discovered' the manifest image much the way Columbus discovered the western hemisphere, but he, like Columbus, was more the explorer than the systematic cartographer, and, as just noted, our Mercator has yet to appear. But what about a Linnaeus, a systematic taxonomist of all the things we find in this new, if all too familiar, world? Ontology, it seems to me, has been resolutely reductive and minimalist in its attempts to come up with an exhaustive list of kingdoms, classes, genera and species of things. No doubt the motivation has been the right one: to prepare the disheveled cornucopia for scientific accounting, with everything put ultimately in terms of atoms and the void, space-time points, or (by somewhat different reductive trajectories) substances and universals, events and properties and relations, As is well known, these procrustean

beds provide metaphysicians with no end of difficult challenges, trying to fit all the candidates into one austere collection of pigeonholes or another. Where do we put numbers, or melodies, or ideas, for instance? Do numbers even exist, and are there different kinds of existence? Instead of attempting any answers of my own to these oft-asked questions of systematic taxonomy, I am going to push, informally, in the opposite direction, like the bestiarists of yore, recounting with as much ontological tolerance as I can, a few of the neglected candidates for thing-hood that everyday folks have no trouble countenancing in their casual ontologies. Metaphysics as folklore studies, you might say.

Quine did some pioneering on this score in *Word and Object* (1960). Are there *sakes*? Are there *miles*? He used these homely examples to loosen up his readers for his theory-driven proposals, his quest for a 'desert landscape' clean-shaven by Occam's Razor, a minimalist ontology for science (and for everything worth talking and thinking about—he was Carnap's student, after all). Since his overall aim was eliminative, he no more had to address the question of whether sakes and miles belonged to different kinds of thing than Linnaeus had to worry about whether mermaids were mammals and dragons were reptiles. Nor were the Oxbridge ordinary language philosophers interested in constructing a systematic taxonomy of the ordinary. Their informal but clever sallies did unearth some useful examples, however, and I want to add some more for the consideration of sophisticated naïve (auto- and hetero-) anthropologists.

Let's consider some of Quine's examples in a little more detail. In addition to Paris and London are there the *miles* that separate them? Are there *sakes*, as suggested by such familiar idioms as 'for Pete's sake' and 'for the sake of — '? What are the identity conditions for sakes? How many sakes might there be? How do they come into existence? Can one still do something for FDR's sake, or did it expire with him decades ago? Art for art's sake shows that not only living things can have sakes. Quine gave us reasons to see the ontological candidacy of sakes as something of an accident of ordinary language (more specifically, of English, since there aren't clear synonyms in all other languages), a sort of potential cognitive illusion frozen in a few idioms, waiting to be vivified by an unwary and profligate semanticist. The restriction to a few idioms ('for the sake of,' 'on behalf of,' and the like) which might rather gracefully be 'fused' into unanalyzed prepositions does suggest, as Quine proposed, that this noun is *defective*, and its putative referents need not be taken seriously from an

ontological point of view. He suggested that we dismiss all the nouns of measure—miles, ounces, minutes, and their ilk—by the same principle, a more demanding exercise of self-denial, since talk about gallons and kilograms and minutes is less easily paraphrased away, and plays a much richer role in daily life.

What about *dints*? Again, the noun, descended from a now extinct term for a blow or strike, has lost its combinatorial power, almost as inert as the *board* in *cupboard*. Not a compelling candidate—on the surface, anyway—for thing-hood, but we shouldn't jump to conclusions. It is noteworthy that computer scientists in the last few decades have found the task of knowledge representation for software agents or bots, where the demands of automated inference put a severe strain on data structures, is often clarified and expedited by the devising of ingenious ontologies that carve up what might be called the manifest image of their bots in novel ways. Dints could conceivably prove to be just the things to represent when reasoning about actions and effects of certain sorts.

Then there are *cahoots* and *smithereens*, both particularly unpromising. These are probably just what Quine said about sakes, degenerate cases that do not require patient elaboration. It is bad enough having graduate students writing whole dissertations on the relationship of constitution between the stone and the statue; we need not create a microdiscipline, smithereen theory, to work on the sub-problem of what to say about all the bits when the wrecking ball hits the statue.

David and Stephanie Lewis (1970) made famous the difficulties of either banishing *holes* altogether or fitting them into any of the austere schemes favored by Quine and others. As they amusingly showed in the dialogue between Argle and Bargle, getting through the day without holes in one's ontology is no easy matter. They never considered the tempting hypothesis that our reliance on the concept of holes has a deep biological or ecological source, since holes are Gibsonian affordances *par excellence*, put to all manner of uses in the course of staying alive and well in a hostile world. Ubiquitous and familiar though they are, it is surprisingly hard to say what holes are made of, if anything, what their identity conditions are, whether they are concrete or abstract, how to count them, and so forth. In the end Bargle says: 'I for one have more trust in common opinion than I do in any philosophical reasoning whatsoever.' To which Argle replies: 'My theories can earn credence by their clarity and economy; and if they disagree a little with common opinion, then common opinion may be

corrected even by a philosopher' (Lewis and Lewis, 1970: 211–12). The sophisticated naïve anthropologist can calmly note the impasse from a position of neutrality and conclude (tentatively) that holes are virtually indispensable in the ontology of the manifest image, but probably not portable, without strikingly counterintuitive revision, to the austere scientific image. But not so fast: holes may play a potent role in organizing the patterns studied in the special sciences—consider the importance of membranes and their multiple varieties of portals in cell biology, for instance—and what about the slit that figures so prominently in quantum physics?

Not so fundamental or indispensable, no doubt, but still not easily dismissed are the *voices* I discussed (Dennett, 1969) at some length at about the same time. My project then, rather like the Lewises' project with holes, was to coax philosophers out of their self-imposed disjunctions about *minds*—are they brains *or* spirits, concrete *or* abstract, *or* are they, strictly speaking, as non-existent as witches and mermaids? I tried to do this by showing that voices, ostensibly nowhere near as mysterious as minds (there is no voice-throat problem to set alongside the mind-body problem, thank goodness), were just as uncategorizable, and none the worse for that. Why are voices more robust candidates for thing-hood than sakes? In part because synonyms or near-synonyms for 'voice' occur in many languages, and the contexts in which these nouns appear are remarkably diverse and varied. You can hear a voice, lose a voice, strain a voice, record a voice, recognize a voice, disguise a voice, and use technology to see through an acoustic disguise to identify the underlying voice, for instance. What *kind* of thing is a voice? It can be strained, like a larynx or knee, but it is not an organ, since it can be lost and recovered. Is it a disposition or prowess or idiosyncratic competence or method, something like a *style*? In a way, but it can also be recorded, like a sound or a melody. Maybe a voice is the sole exemplar of an ontological kind all of its own. Maybe bass voices, soprano voices, childish voices, raspy voices and nasal voices are all kinds of voices but voices aren't kinds of anything at all except the voice kind. (There could be other such kinds, of course. In some cultures that use bells a lot, there might be the recognizable *ding* of this or that bell.) Here is an interesting question in the empirical theory of perception and knowledge as it merges with semantics (see Jackendoff, 2002, 2007): why is it so natural for human minds speaking many different languages to hit upon voices as a valuable addition to their ontology? But remember: we

are bracketing for the time being the question of which elements of the ontology of the manifest image we take ourselves to be committed to.

Haircuts (and hairdos more generally) are another variety that is strangely resistant to pigeonholing. A haircut is not really an event or a property or a relation or a material object (made of hair, presumably). How long haircuts last, and just how to apply the type-token distinction to them (think of the Beatles haircuts all around the world), are not readily answered questions.

The wonderful world of economics provides us with further socially constructed things, such as dollars and euros. Do dollars exist? It seems obvious that they do, and—unlike some of our other problematic things, they seem eminently countable, designed to be countable in fact. But we should note, before moving on, that there has been a historical progression in the direction of abstraction. No doubt a large source of the confidence with which we countenance the existence of dollars as things is that today's dollars are the descendants, in effect, of paradigmatic things, made out of metal and having a shape and weight, like dimes and nickels and silver dollars. These were followed by first paper dollar *bills*, five-pound *notes*, and their kin, and then entirely abstract dollars, by the trillions. It's a good thing we don't have to make them all out of metal; they would bury the world. We gain a tremendous efficiency from this transition, and all we lose is the standard way of keeping track of individual concrete things. It makes sense to ask whether this particular dime or silver dollar once passed through the hands of Jacqueline Kennedy, but not whether the hundred dollars I just sent to my PayPal account ever did.

The concept of abstract dollars is undeniably useful in our everyday affairs, but is utility tantamount to reality? This question sits happily within the perspective of sophisticated naïve anthropology but is out of place in traditional metaphysics. Consider the reality of centers of gravity or mass. Are they real or not? Some philosophers have said yes, and others no, and both sides have used their conviction to disagree with my use of them as a useful analogy for the soul, as the 'center of narrative gravity' (Dennett, 1991). Compare the concept of the center of gravity of, say, an automobile with the concept of Dennett's lost sock center, defined as the center of the smallest sphere that can be inscribed around all the socks I have ever lost in my life. The former is an undeniably useful concept, and so *palpably* useful that you can almost feel a center of gravity once you know what to feel for. The toddler's toys, called Weebles, 'wobble but they don't fall down';

watching a child discover this marvelous property by exploratory manipulation, it is hard to resist casting the task of *finding the center of gravity* alongside the task of *finding the hidden coin* as different instances of the same activity. But however perceptible or tangible a center of gravity is, it is (one would suppose, as a metaphysician) in exactly the same basic ontological genre as Dennett's lost sock center, and kazillions of other equally pointless abstractions we could invent. Must we fill our world with all that pseudo-Platonic garbage? Scientific utility, as Quine never tired of reminding us, is as good a touchstone of reality as any, but why shouldn't utility *within the manifest image* count as well? Is there anything dangerously relativistic in acknowledging that the two images may have their own 'best' ontologies, which cannot be put into graceful registration with each other? (This suggestion is a close kin, and descendant, of Quine's ontological relativity (1969), and his claims about the indeterminacy of radical translation, but I cannot do justice to exploring the relations in the limited space of this essay.)

In the introduction to *Brainstorms* (1978), I imagined a society that speaks a language just like English except for

> one curious family of idioms. When they are tired they talk of being beset by *fatigues*, of having mental fatigues, muscular fatigues, fatigues in the eyes and fatigues of the spirit . . . When we encounter them and tell them of our science, they want to know *what fatigues are* . . . what should we tell them? (pp. xix–xx)

Our task, I suggested then, and still maintain, is more a matter of *diplomacy* than philosophy. It is not that there is or must be—there *might* be—a univocal, all-in, metaphysical *truth* about what there is, and whether fatigues can be identified as anything in that ultimate ontology, but just better and worse ways of helping people move between different ontological frameworks, appreciating at least the main complexities of the failures of registration that are encountered. For anyone interested in such diplomatic projects, a vigorous anthropological scouting is good preparation.

Metaphysicians and philosophers of mathematics have devoted years of work to the question of whether numbers exist, while taking the existence of *numerals* more or less for granted, but why? Numerals, as tokens, are no doubt unproblematic trails of ink or chalk or some other physical particular, but as types, they are more problematic. They are like words in this regard. Words are rather like voices: they seem to be a set with no obvious

superset. Nouns and verbs are kinds of words, but what are words a kind of? Symbols? Sounds? I have suggested that words are a kind of *meme*; words are memes that can be pronounced (Dennett, 2009). They are also, as Jackendoff (2002) puts it, 'autonomous, semi-independent information structures, with multiple roles in cognition.' In other words, they are a kind of software that runs on the brain.

And how about software? Does it exist? I have been fascinated to learn recently that there are philosophers who are reluctant to countenance software (and all its subspecies, from data structures and subroutines to apps and widgets) as being among the things that truly exist. This is a descendant, apparently, of a similar line of ontological skepticism or mistrust about information in general, as something—some thing—that can be stored, moved, processed.

Once again, the perspective I would recommend is that of the diplomatic anthropologist, not the metaphysician intent on limning the ultimate structure of reality. The ontology of everyday life is now teeming with items that, like fatigues, sit rather awkwardly in the world of atoms and molecules. If we can understand how this population explosion came about, and why it is so valuable to us as agents in the world, we can perhaps discharge our philosophical obligations without ever answering the ultimate ontological question. To me it looks more and more like professional make-work, an artifact of our reasonable but ultimately optional desire for systematicity rather than a deeper mystery in need of solving.

References

Bobrow, D. (1985). *Qualitative Reasoning about Physical Systems*, Cambridge, MA: MIT Press.

Dennett, D. (1969). *Content and Consciousness*. London: Routledge & Kegan Paul.

—— (1978). *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery, VT: Bradford Books.

—— (1991). Real Patterns. *Journal of Philosophy*, LXXXVIII: 27–51.

—— (2005). *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge, MA: MIT Press.

—— (2006). Higher-order truths about chmess. *Topoi*, 25: 1–2.

—— (2009). The cultural evolution of words and other thinking tools. Cold Spring Harbor Symposia on Quantitative Biology, 74: 435–41 [online] <http://ase.tufts.edu/cogstud/papers/coldspring.pdf> accessed 11 May 2012.

Hayes, P. (1978). The naïve physics manifesto. In D. Michie (ed.) *Expert Systems in the Mocroelectronic Age*. Edinburgh: Edinburgh University Press.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

—— (2007). *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge, MA: MIT Press.

Ladyman, J. and Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Lewis, D. and Lewis, S. (1970). Holes. *Australasian Journal of Philosophy*, 48, 2: 206–12.

Quine, W. (1960). *Word and Object*. Cambridge, MA: MIT Press.

—— (1969). *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Sellars, W. (1963). *Science, Perception and Reality*. London: Routlege & Kegan Paul.

6

The World in the Data

James Ladyman and Don Ross

1. Science and fundamental ontology

Ladyman and Ross (2007) present a version of naturalized metaphysics to weakly unify the special sciences by reference to fundamental physics. 'Fundamental' physics here means that part of physics that is implicitly tested, or could explicitly be confirmed, by every measurement that could be taken at any scale of reality and in any region of the universe. The unification is 'weak' because it is not reductionist. Fundamental physics constrains all special sciences (including the parts of physics that are not fundamental), but it doesn't track information at the multiple scales necessary for capturing most of the 'real patterns' (Dennett, 1991) that are the objects of study in chemistry, biology, geology, economics, and so on. Ontology is scale relative: all real patterns except those of fundamental physics are detectable at some scales but not at others. The same point applies to everyday phenomena tracked by reference to the parochial practical purposes of people, such as tables, hockey games, and capital cities. Mere patterns—stable but nonredundant relationships in data—are distinguished from 'real' patterns by appeal to mathematical information theory. A pattern is redundant, and not an ultimately sound object of scientific generalization or naturalized ontology, if it is generated by a pattern of greater computational power (lower logical depth).[1] Then to be is to be a real pattern. Ladyman and Ross provide reconstructions of important concepts in philosophy of

[1] The allusion here is to Charles Bennett's (1990) idea of logical depth, which he proposes as a measure of complexity. The idea is roughly that complex objects have structures that are in each instance probably produced by long causal histories. A problem with Bennett's definition for practical purposes is that it is not computable for a given data set (Ladyman,

science such as causation and laws in terms of recurrent types of structural relations among real patterns. Most important claims that have been thought to be laws in the history of modern science describe such structural relations in mathematical terms that survive episodes of theory change in approximate form. Individual objects as used by people for coordinating reference to a universe organized from specific parochial perspectives are real patterns of relatively high logical depth and thus do not feature in scientific generalizations. Furthermore, the important real patterns in science are not reducible to facts about the intrinsic properties or natures of individual objects. Ladyman and Ross defend a metaphysics that does not take individual things to be fundamental (Ladyman, 1998 and 2002).

If metaphysics is to be part of the pursuit of objective knowledge, it must be integrated with science. Genuinely naturalized metaphysics must go beyond mere consistency with current science; it must be directly motivated by and in the service of science. The linked collection of institutions that constitutes science—its academic departments, research institutes, journals, and granting agencies—are the only ones dedicated to self-correcting pursuit of objective knowledge as their primary motivational goal. Many other institutions of course gather knowledge; but only in science is it a decisive objection to a practice if one finds that it makes knowledge accumulation subordinate to another goal.

Ladyman and Ross (2007) argue that making the world intelligible and comfortable using humans' intuitive conceptual categories, which is arguably the primary goal of most *non*-naturalized metaphysics, is not only a different goal from that of contributing to objective knowledge, but also a systematically conflicting one. Van Fraassen (2002) agrees; a metaphysics that is not at least broadly true, he maintains, is worthless. Van Fraassen also believes that metaphysical truths cannot be aimed at with reasonable prospects of success—perhaps they could be obtained by luck, but then we wouldn't recognize our good fortune—and so he concludes that metaphysics should be abandoned. The naturalistic metaphysician, by contrast, is optimistic about the possibility of bringing metaphysical hypotheses into closer conformity with objective reality to the extent that these hypotheses non-trivially unify bodies of established scientific knowledge.

Lambert, and Weisner, forthcoming); but this issue doesn't impugn the theoretical use to which we put it.

Many scientists rhetorically share van Fraassen's pessimism, and express it in categorical terms when they publish reflections on the broadest explanatory ambitions of science. Stephen Weinberg (1992) argues that philosophy—by which he means both metaphysics and epistemology—is always unhelpful to science, and often impedes it outright by temporarily erecting barriers to progress through conservative allegiance to well-established concepts. Hawking and Mlodinow (2010) pronounce philosophical efforts to expound on the grand structure of reality 'dead' and claim that physical cosmology has entirely supplanted philosophy's former role. The meaning of these assertions is ambiguous, however. Weinberg and Hawking and Mlodinow in fact concede that unification of science to create comprehensive large-scale models of reality as a whole is a valued scientific objective—if not indeed the supreme such objective—but think that its institutional home has shifted from departments and journals called 'philosophy' to departments and journals called 'physics'. They should thus be counted as among the sympathizers with naturalized metaphysics, their anti-metaphysical rhetoric notwithstanding.

Most popular naturalistic metaphysics produced *directly* by scientists is of low quality. In the case of Hawking and Mlodinow, it is egregiously sloppy. Even if, as they maintain, the generalization of string theory by M theory unifies theoretical physics,[2] it makes no evident promise of unifying the sciences in general, so could hardly supplant the whole of metaphysics. Naturalistic metaphysicians may be disinclined to waste critical attention on such shallow efforts, and, if they are philosophers, the incentives of the academy support such fastidiousness. For good reason, work mainly aimed at popularizing knowledge is not a proper part of the scientific literature, but at best a second-order description of that literature.

Ladyman and Ross placed their 2007 brief for and outline of the contents of naturalistic metaphysics firmly within the polemical environment of philosophy. Their almost complete lack of shared motivations or fundamental premises with non-naturalistic (analytic or, pejoratively, 'neo-scholastic') metaphysicians precluded constructive engagement with that part of the philosophical literature, though they did not shrink from destructive challenges to it. But they remained firmly inside the mansion of institutional philosophy by anchoring their arguments closely to recent

[2] Following Smolin (2007) and others, we would bet against this speculation; and in any event speculation, running far ahead of empirical confirmation, it assuredly is.

themes in philosophy of science that have avowed or implicit metaphysical importance.

The discourse should not remain cloistered in this way, however. Naturalized metaphysics, if it is to be a genuine contribution to knowledge, should not mainly consist of borrowing from science to serve dialectics internal to institutional philosophy. Talk of naturalizing metaphysics is little more than hopeful rhetoric if the enterprise does not aim to play a reflexive role in guiding the formulation and elaboration of hypotheses that scientists actually investigate. Such positive influence on knowledge is unlikely if, for the sake of connection with literature most scientists have not read, emphasis on themes such as the nature of causation crowds out interventions in specific theses promoted by scientists.

In this chapter, therefore, we reformulate the main conclusion of Ladyman and Ross (2007) in terms intended to provide better tools for application to scientific controversies. We build our rhetorical bridge by beginning with a famous philosophical slogan of Wittgenstein's that many scientists will know, and of which some will approve. In the *Tractatus Logico-Philosophicus* (Wittgenstein, 1922: 5) Wittgenstein said that 'The world is the totality of facts, not of things.' This claim has doctrinal aspects in common with our view, specifically in calling attention to the ontological significance of structure; but the structure in question is logical rather than mathematical structure, and Wittgenstein certainly doesn't deny that things are basic constituents of reality. We here adopt the form of address and twist it as follows to state our metaphysical thesis: the world is the totality of non-redundant statistics (not of things). The aim of the chapter as a whole is to say what we mean by this, and why we believe it.

The reader might reasonably doubt that this formulation should have any more resonance with scientific argument than the more abstract and dialectical presentation in Ladyman and Ross (2007). However, we will show that the issues are not parochial to philosophers by taking as our principal argumentive foil a recent effort by a leading theoretical physicist, David Deutsch (2011), to state a comprehensive philosophy of science, including strong metaphysical theses in our sense, motivated by unifying themes drawn mainly from three sciences: quantum physics, computation theory, and evolutionary theory. We will not attempt either a full exposition or a thorough assessment of Deutsch's views. Rather, we use Deutsch's philosophical opinions as a vehicle for doing two things. First, we aim to demonstrate the relevance of naturalistic metaphysics to some

current real controversies within science, as identified by an important scientist. Second, we aim to show how metaphysical considerations can inform scientific debate without becoming preoccupied with analysing concepts or language. This preoccupation is profoundly at odds with the point of naturalism, and is responsible for the barely disguised, or undisguised, hostility with which many scientists regard philosophers.

Deutsch, like Ladyman and Ross (2007), emphasizes the ways in which science uniquely opens the way to *limitless* expansion of knowledge by disregarding parochial human intuitions about what 'makes conceptual sense'. Innovations in science recurrently confound these intuitions, leading cultural expectations about what is 'reasonable' to adapt and adjust. Most philosophical commentary on science that is rooted in conceptual analysis amounts, even if inadvertently, to efforts at domesticating science and ultimately restricting its reach. Ultimately, we will criticize Deutsch for imposing an unwarranted restriction on his own understanding of the resources with which scientific activity engages with the general structure of reality. In particular, Deutsch cannot accept the possibility expressed by our pseudo-Wittgensteinian slogan above, and this in turn leads him to close off live options in fundamental physics through what, in our terms, amounts to a priori conservatism. In this he follows the consensus in the contemporary philosophy of physics literature, according to which one of several alternatives to Bohr's original version of the Copenhagen interpretation of quantum mechanics must be favoured on the grounds that Bohr's picture is so 'incomplete' as to be worthless. This situation reminds us that maintaining a truly consistent naturalistic attitude requires critical vigilance. Two thousand years of inherited philosophy tugs on our ankles, our cognitive inertial mass being our evolutionarily endowed perceptual systems with their attendant higher-level cognitive structure of physical objects, animals, and plants from which the philosophical system of Aristotle was originally abstracted.[3]

The chapter proceeds as follows: in the next section we review relevant aspects of Deutsch's philosophy; in section 3 we discuss standard realism

[3] Another physicist who is led by his science into metaphysics and who also reaches a similar conclusion to us about its value when done through a priori analysis is Bernard d'Espagnat, who complains that science has shown that the 'very notions that Aristotle and his followers got us into the habit of considering to be fully primary and basic' are in fact anthropocentric (2006:1). We shall return to his views below.

and instrumentalism, and introduce Deutsch's argument for incorporating Everett's interpretation of quantum mechanics into metaphysics; in section 4 we discuss the metaphysics of quantum mechanics; and finally in section 5 we present Peirce's naturalistic embrace of the probabilistic modal structure he found in some science, and defend the claim that the world is fundamentally statistical.

2. A physicist's philosophy

Deutsch's *The Beginning of Infinity* (2011) is resolutely and inspiringly optimistic about the potential for growth in scientific knowledge, and in consequence about the capacity of people to transform themselves and their social and physical environment for the better. Because the potential for new knowledge is limitless—so long as we protect and enable the institutions of science—we never have a basis for regarding any problem as insoluble. We endorse this confidence in the values of the Enlightenment, and applaud Deutsch's rousing denunciations of timid and repressive anti-scientific conservatism. Philosophical analysis cannot be expected to furnish a significant resolution to this battle between fundamentally conflicting attitudes. However, scholarship in a much more general sense is highly relevant to the fight, since both scientism and conservatism are in the final analysis interpretations of history. Furthermore, as argued in the opening and closing pages of Ladyman and Ross (2007), the best motivation for trying to synthesize our scientific knowledge into a unified picture—that is, for building naturalistic metaphysics—is the crucial service this activity potentially performs in extending the Enlightenment project. If science is not seen to provide the basis for a general worldview, then people will continue to collectively confabulate alternative general pictures. This in turn matters because the confabulated pictures inspire groundless and usually wasteful and destructive politics and policy. We see no reason to be coy about the fact that, like the logical positivists, our philosophizing is inspired by a normative commitment: while acknowledging the importance of conserving what is valuable, we abhor conservatism, which we view as a sad refusal to explore the magnificent range of possibilities that our ability to do mathematics allows us, and thus betrays the best reason for caring passionately about objective truth.

The best general philosophy of science reveals a difficult tension between the ideas of objectivity, on the one hand, and readiness to let

the flow of new knowledge smash through conservative conceptual barriers, on the other hand. The tension in question is nicely demonstrated in Michael Friedman's efforts to extend Kant's philosophy of science in a strongly naturalist way, a project Friedman has shown through his historical scholarship (Friedman, 1999) to have been the basic objective of our shared logical positivist heroes. The ambition is to safeguard the objectivity of scientific theory choice without downplaying the significance of the kind of radical theory change that has repeatedly swept through science, and especially through physics, since Kant sought to ground the objectivity of Newtonian mechanics in synthetic a priori knowledge of necessary truths. Friedman (2001) rehabilitates the positivist project of separating the constitutive role of the a priori in establishing the framework within which empirical testing is possible from the idea that the content of the a priori is necessary and not revisable. For example, as discussed by Friedman in this volume, the concept of an inertial frame exemplifies the general fact that theories cannot be tested without background assumptions that implicitly define certain concepts needed to connect theoretical structure (in this case the notions of force and mass) with empirical data (such as the paths of the planets across the sky). Newton's first law only holds in inertial frames in which, as a conceptual requirement, force-free bodies travel along straight lines. However, one can only establish that a body is force-free via Newton's second law, which also only holds in inertial frames. The circularity here, while not vicious, has the consequence that the definition of an inertial frame cannot be tested directly but is rather required to get the business of testing Newton's laws off the ground. Yet this status did not prevent Einstein from fundamentally altering the idea of an inertial frame in General Relativity. Thus, Friedman concludes, constitutive a priori principles can be revised and are not necessary; but while we have them they provide objective structure to scientific inquiry that avoids Quinean holism about confirmation and the attendant problem that theory choice seems to become a matter of taste (Quine, 1969).

Objectivity grounded in concept specification may not appear to be the real thing to many scientific realists—and certainly not to Deutsch. However, let us put this purely philosophical issue to one side for now. Our concern here is that Friedman's picture enjoins the same kind of cognitive inertia that Kuhn (1962) attributes to normal science. If Friedman's account is correct, then unless a scientist is ready to replace the whole constitutive framework she must accept the a priori content of existing

science, and, since there are lots of incentives against doing the former, cognitive conservatism follows. Of course there *is* much cognitive conservatism in science, enforced by its peer review institutions. But to elevate conservatism to be part of a deep *logic* of science welcomes what Kuhn called 'the essential tension' between science as juggernaut and science as custodian of accepted wisdom. Friedman's work reconciles the positivists with Kuhn. But the picture of science as, at any given time, either in hypercautious normal mode *or* in ferment and riot mode—Kuhn's revolutionary episodes—where rationality gives way to idiosyncrasies of taste is hardly comforting for those who hope that science might enable us to get a steadily clearer grip on the structure of general objective reality. Even those who consider Friedman's account to be accurate with respect to normal science should recognize that it lacks resources for understanding the preservation of objectivity through episodes of profound theory change.

The leading critic of the positivists in their heyday was of course Popper. There has been a tendency for philosophers of science to regard Popper as something of an embarrassment. However, naturalistic philosophers should take it as an interesting fact about science that Popper has long been the favourite philosopher of science among scientists; and it would be condescending to attribute this entirely to the fact that Popper's philosophy is relatively non-complex, self-contained, and flattering to scientists.[4] A necessary condition for Popper's appreciation by scientists is that he emphasizes themes they agree to be central: the provisional and dynamic nature of their knowledge, recognition of the creativity and indeed artistry involved in hypothesis formation, and acknowledgement that commitment to the institutional norms of scientific practice is an ethical stance with sweeping social and political implications. None of these themes are unique to Popper; but in his work they are fused into a philosophy that may not provide scientists with what philosophers wish scientists would want from them, namely metaphysical truth, but instead provide scientists with something else they value: affirming inspiration. Deutsch is an avowed and uncritical follower of Popper, as one might

[4] We have in mind here the fact that Popper often writes as if individual scientists, showing moral selflessness unique among modern institutional roles, heroically strive to refute their *own* theories. It is often pointed out that this describes the actual activity of very few scientists. However, it is easy to strip this silliness out of Popper's picture by noting that the greatest rewards in science are bestowed on those who make major revisions in the theories of others.

expect, seeing himself as reaffirming and extending the latter's neo-Enlightenment manifesto.

We fully share this normative-institutional commitment. In welcoming much of what Deutsch has to say about science, we will thus implicitly be celebrating Popper too. We will ultimately concur with the majority of philosophers that *metaphysical* inspiration is not to be found in that source, and we will criticize Popper's and Deutsch's dismissal of the positivists. Our aim is to show that when Deutsch's errant metaphysical opinions are corrected, the result is a picture of science in which ontological *and* progressive enlightenment are mutually complementary.

One of Deutsch's leading themes is *not* taken from Popper. This is his repeated claim that the most important epistemic function served by a scientific theory is not prediction or formal generalization, but explanation. Unfortunately he is not very explicit about what he understands explanations to be. He says only that they are 'assertions about what is out there and how it behaves' (Deutsch, 2011: 3). This will strike philosophers as comically imprecise. However, Deutsch's subsequent extensive discussions of examples makes clear that he means by explanation what most philosophers of science do: the embedding of claims about facts within networks of causal and nomological generalizations. Such explanations, he argues, here following Popper, are creative conjectures that resist easy adjustment in the face of recalcitrant tests. The epistemological ideal of falsifiable explanations that survive efforts at actual falsification is supplemented by a complementary principle of maximum explanatory 'reach'. This is a Lakatosian virtue: it is the capacity of explanatory strategies—which Deutsch roughly identifies with the scientific theories that form the cores of Lakatosian research programmes—'to solve problems beyond those that they were created to solve'. 'Reach', then, is what philosophers of science refer to as the capacity of theories to produce novel predictive success, and to support consilience and unification.

Deutsch emphasizes, just as Ladyman and Ross (2007) do, the remarkable fact that theories designed to explain phenomena observed locally and at parochial length and time scales familiar to humans have so often proved applicable to remote regions, and at more general scales than any human can directly experience. So, for example, we have never been anywhere near stars or their cores yet our theories explain their behaviour and the phenomena that they cause closer to home (p. 3). Deutsch emphasizes, however, that reliance on examples of this kind can mislead us into

thinking that how things seem to us pre-scientifically, on the basis of our limited engagement with our local environment and interests, should automatically generalize to less anthropocentric scales. He defends a principle of anti-parochialism, according to which we should self-consciously guard against mistaking appearance for reality and confusing local regularities with universal laws (p. 76). Furthermore, he argues that science and philosophy should take care to avoid a more specific sense of anthropocentrism, the projection of features of our experience and cognition on to the non-human world. On the other hand, he argues that the 'Principle of Mediocrity' (p. 76), according to which humans are not cosmically special or significant, is also false, on the grounds that in certain respects, namely our capacity to generate knowledge and understand abstractions, and to radically redesign our own environments, humans are as far as we know unique among physical entities. Ladyman and Ross (2007: 2) stress our ability to do mathematics as a necessary condition for this. Deutsch often writes as if that ability is also a sufficient condition for the limitless proliferation of explanations with reach, so long as scientific institutions are not politically repressed. We think this idea is plausible, and we hope the future bears it out.

Deutsch grounds these principles in an account of observation and the evidential basis of science. He argues against 'empiricism', defined as the view that all knowledge is inductively derived from sensory experience. Again following Popper, Deutsch asserts that explanations cannot be based on induction because they transcend the features of the observations they explain—and not merely logically, in the sense of universally quantifying over them, but qualitatively. A second aspect of Deutsch's anti-empiricism is his insistence that scientific observation is inherently theory laden, in the sense that all data are gathered, transduced, and systematized by means of instruments built to purposes informed by theory. Deutsch does not take this to impugn the status of observation, after the fashion among strong social constructivists, but rather to enrich the range of what ought to count as instances of observation. So, for example, an expert eye studying a photographic plate exposed by a telescope sees stars. Among candidates for the status of 'genuinely observed' in the processing chain leading back from worldly objects, to images on the retina, to electrical signals in the optic nerve and so on, the project of selecting a stable 'winner' by means of philosophical analysis is misguided in principle. Thus we should conclude that if our theory tells us that what produced a visual image used to

represent the data from a radio telescope is a galaxy then we are entitled to say that an astronomer looking at the image sees the galaxy.

According to Deutsch, empiricism historically replaced a conception of knowledge based on human authority with one based on the equally false authority of sense experience and induction. Endorsing the view that there is no 'reliable means of justifying ideas as being true or probable', he favourably contrasts this 'fallibilism' with 'justificationism'.

Here we find Deutsch embracing an egregiously uncharitable Popperian criticism of the positivists, mixed with sound insight into the theme identified at the opening of the present section that goes to the core of the relationships between naturalism and Enlightenment values. The confusion is that Deutsch clearly imagines—perhaps misled by the surface semantics of the word 'confirmation'—that the aim of gathering evidence for scientific generalizations is *certainty*. Deutsch rightly castigates such an aim as contrary to the scientific attitude; however its attribution is utterly unfair to his leading examplars of 'bad philosophy', the logical positivists and empiricists. Early in his book he denounces justificationism for embracing the project of 'securing ideas against change' (p. 9). As discussed earlier, however, the positivists' fundamental project was to reconcile Kant's project of grounding the objectivity of specific observations in prior categorical and intellectual structure with the recognition, forced by Einstein's physical vindication of non-Euclidean geometry, that apparently a priori grounds are revisable in light of experience. As explained by Friedman (1999), the positivists resorted to conventionalism as a way of achieving this reconciliation. This offers something weaker than what Deutsch's ultimately ontological quest for objectivity demands. However, since conventions are obviously provisional by nature on anyone's account, it is clearly rash to accuse the positivists of hoping to *block* conceptual or theoretical change. On the other hand, in supposing that our confidence in science might be buttressed by appeal to extra-scientific elements of thought and belief—specifically, in the case of the positivists, to formal logic—the positivists indeed accepted the anti-naturalist view of scientific beliefs as standing to benefit from extra-scientific securitization. For Carnap, Neurath, or Reichenbach the securitization in question was not so much against change as against instantaneous doubt. However, early analytic philosophy indeed thereby opened the door, in its founding documents, to what has since metastasized, in contemporary analytic metaphysics, into a whole-scale elaboration of scientifically untutored

and logically elaborated intuitions, often for purposes of defence against challenges to conceptual comfort from science (Ladyman and Ross, 2007, Chapter 1).

We do *not* here mean to imply that if we could bring the positivists into the twenty-first century they would be sympathetic to analytic metaphysics. Friedman has shown us exactly what they could think instead—how, indeed, they could comfortably maintain their naturalism.[5] Our point, instead, following Wilson (2006, and this volume) is that in assigning a foundational role to *logic* as a source of 'clarification' in specifying the meaning of scientific propositions—and thereby 'clarifying' scientists' ontological commitments—the positivists in all innocence prepared the boggy ground in which so much philosophy flounders today, where instrumental control of the world of experience is granted as the domain of science and 'deep' ontology is assigned as a task for a priori analysis.

Deutsch has no patience with instrumentalism, defined as the view that science merely summarizes observations while avoiding association of theoretical structures with commitment to an underlying reality that corresponds to them (p. 31). He instead advocates realism, defined as the view that we can and do have knowledge of the actual structure of the physical world. He argues that science is progressive in the sense that the truth of old theories lives on in their successors (p. 113), maintains that this applies not only to predictions but also to explanations, and so asserts that events like the replacement of Newtonian mechanics by quantum mechanics is no threat to realism. Ladyman and Ross (2007) also defend all of these theses, based entirely on premises about the history and practice of theory construction and adjustment in the history of physics. Finally, Deutsch's realism incorporates denial of the empiricist idea that scientific theories can generally be separated into formal cores and empirical interpretations. This is an interesting and controversial thesis that can only profitably be argued on the basis of carefully considered examples from science. It has received a sustained and highly persuasive such defence in Wilson (2006) (see also Wilson's second chapter, Chapter 9, in this volume). We will return to the issue in some detail below.

The claims reviewed to this point are the basis for Deutsch's positive philosophy of science, including its naturalized metaphysics. This consists

[5] Historians of philosophy argue over the extent to which the actual Carnap anticipated or indeed fathered a different version of naturalism, namely Quine's—which, as Friedman complains, in many moods suggests giving up on objectivity.

of two fundamental components. The first is a repudiation of reductionism, and a commitment to the reality of causally efficacious emergent phenomena. The second is the claim that abstract entities can be real and can have real causal powers, notwithstanding that causation is itself held to be an abstraction (p. 124). Deutsch argues for the first claim just as Ladyman and Ross (2007, Chapter 4) do, by invoking the fact that in scientific practice, explanations of macro-scale phenomena in microphysical terms are almost never available. Rather, micro-scale complexity resolves itself into macro-scale simplicity (pp. 108–9), and so we explain macro-scale phenomena in macro-scale terms. Macro-scale entities, properties, and processes are emergent but since they figure in explanations and indeed causal explanations they are nonetheless real. By the same token, since sometimes the best explanation of a phenomenon cites an abstract entity or property, we ought also to believe in the reality of such abstractions. For example, the best explanation of a physical event such as the removal of a chess piece from a board may be that a computer programme is better than its human opponent at chess (pp. 114–15). As Ladyman and Ross (2007), along with Batterman (2002) argue at length, this is the only picture that allows us to recognize that sciences other than fundamental physics can and do discover first-order knowledge, and not merely approximate second-order gerrymanderings of physical phenomena that happen to impinge on our parochial interests and projects.

Deutsch's commitment to the reality of abstractions has two important special cases: universal functionality and representational capacity in digital systems of abstractions such as alphabets and the genetic code (Chapter 6); and, the reality of computation of information where the latter is construed as 'a physical process that instantiates the properties of some abstract entity' (p. 194). Since he believes that human brains effectively instantiate universal Turing machines, Deutsch thinks of human intelligence as limitless in the sense of implementing a universal capacity for generating explanations. The restriction of universality to digital systems arises from the fact that only in such systems should one expect a monotonic relationship between the correction of specific errors and improved overall structural correspondence with reality. The objective reality of information and computation has an important consequence in Deutsch's discussion of quantum mechanics, where it is the basis of one of his arguments for the reality of the multiverses featured in the Everett interpretation.

Deutsch is adamant that his denial of reductionism does not lead to holism. Holists, he argues (p. 110), accept that reductionism correctly describes scientific reasoning patterns but combine this with a (correct) metaphysical conviction that wholes often have both structural and causal properties shared by none of their parts, and therefore tend to reject the suggestion that science can furnish an adequate general characterization of the structure of reality. The denial of holism is also crucial to Deutsch's understanding of the reality of Everettian multiverses. According to him, what physicists—and indeed all scientists who achieve explanatory success—describe are *histories* of causal regularities that are projectible (that is, stable enough to explain and control) because they are relatively 'autonomous'. By this Deutsch means what recent philosophers such as Cartwright (1983, 1989) and Mäki (1992) intend by speaking of relative 'isolation'—they function as machines that, suitably abstractly characterized, transmit information. Everything about which scientific generalizations can be had—or, to use Deutsch's own language more exactly, everything that can be a target of explanation, which is to say everything real—is such a history.

Thus, for Deutsch, to be is to be 'an emergent quasi-autonomous flow of information in the multiverse' (p. 305). Minus the commitment to the multiverse as the interpretation of the fundamental structure of reality, this corresponds closely, though not exactly, with what Ladyman and Ross (2007, Chapter 4) intend when they summarize their positive naturalized metaphysics by means of the slogan that 'to be is to be a real pattern'. Because, like Deutsch, Ladyman and Ross (2007) assume that the pursuit of knowledge is without limit, they also share his view that we can never be sure that anything we now regard as a real pattern will be so regarded in future; on the other hand, we draw confidence from the historical fact that the history of science is one of continuous refinement of physically applied mathematical structure, rather than abandonment and replacement. If reality were unifiable by reduction, it would follow from this conception that in the limit there is only one real pattern, perhaps resembling the 'blobject' of Horgan and Potrc (2008) in some respects. However, the denial of reductionism allows for the multiplicity of real patterns recognized in Ladyman and Ross's 'scale relativity of ontology'.

The information-theoretic conception of existence that Deutsch shares with Ladyman and Ross (2007), as described above, is compatible with the possibility that there is no general, over-arching account of the relationships among the real patterns. That is to say, the possibility remains that the

world is fundamentally disunified, as urged by such philosophers as Dupré (1993) and Cartwright (1999). The investigation of the question of unifiability, based on—and only on—the growing body of scientific knowledge, and especially on knowledge of fundamental physics, is precisely the project of naturalized metaphysics. Like all objective inquiry, it is fallible: any current naturalistic metaphysics is almost certainly wrong in some respects, and might be wrong in general; or evidence for disunity might begin to mount, and we would come to suspect that metaphysics has no positive content beyond the abstract characterization of all that exists as real patterns.

We end this partial account of Deutsch's views, and comparison of them with other perspectives of immediate interest, with the observation that in exact accordance with Ladyman and Ross's (2007) defence of naturalized metaphysics, what motivates this particular scientist to write a work of philosophy is his sense of the value of a unified account of physics and the special sciences, especially those concerned with cognition, language, and evolution. This lends support to our view of the point of metaphysics in a scientific age. If we had good reason to believe that the disunity theorists were right then there would be little point to Deutsch's endeavour, since he adds nothing to the content of any particular science, but seeks only to say something about how the sciences as a whole hang together, and to defend the metaphysics of the Everettian multiverse in the light of the former account. Furthermore, his project also lends some support to our claim that the scientist seeking insights in pursuit of unification will find nothing of any use in most of the literature produced by analytic metaphysicians. Deutsch is interested in understanding the world in the light of the fact that we have evolutionary psychology and linguistics, genetics, cosmology, computer science, and quantum mechanics. Contemporary discussions of whether statues are identical with lumps of clay, the special composition question in mereology, presentism versus eternalism, and so on are completely disconnected from his concerns because those debates utilize no information about the world that has been learned from any sciences. They may as well be occurring in regions of space-time that are space-like separated from those in which we and Deutsch sit; we recognize nothing in them that corresponds to any reality we can measure.

However, we have also suggested that in slavishly following Popper's critique of the positivists, Deutsch makes no progress with respect to the tension between the pursuit of objectivity and the struggle against

conservatism. Indeed, we will argue that in this respect he slides backwards and falls into dogmatism at the level of *physics*. But our positive alternative to Deutsch's view will not defend the positivists as Friedman does. We don't follow that route because, as argued earlier, Friedman's account is unsatisfactory with respect to theory change. An aspect of positivist thought that Friedman does not exploit and which divides the positivists from Kant is precisely the issue on which Popper leads Deutsch most damagingly astray. The positivists, notably Carnap and Reichenbach, understood that induction does not lead to certainty, while recognizing its centrality to scientific knowledge. They consequently played major roles in the development of modern statistical inference. Deutsch is blind to the fact that there are two distinctions to be made: one between those who think there can be a positive evidential support relation between data and theories, and those who don't, and the other between those who think that one can be justified only if one is certain (infallibilists) and those who think that one can be justified and uncertain. Popper happened to be a fallibilist anti-inductivist; but it is also possible to be a fallibilist inductivist, as the positivists were. They saw that Kant was wrong about more than just the necessary physical truth of Newtonian physics and Euclidean geometry; he was also wrong to think that apodictic reasoning was the only variety of justification, and that causation and law could analogously only be deterministic and necessary and sufficient for their effects. As we will see, Deutsch somewhat ironically repeats the second error of Kant's, and thereby falls into a conservative trap himself.

We have seen that Deutsch is committed to the idea that science, through its drive for unification, licenses and demonstrates the value of a naturalistic metaphysics. Through criticizing him, we will demonstrate the basis for a metaphysical picture that makes better sense of the science.

3. Realism, instrumentalism, and quantum physics

To inquire into realism about quantum physics is inevitably to be drawn to at least some extent into the debate between realists and anti-realists as conducted by philosophers of science *and* physicists over several decades. We will not review the history of that debate here—our views are explained in Ladyman and Ross, 2007, Chapter 2—but will directly

contrast the positive position we defend with Deutsch's view. The latter is particularly well-suited to our polemical purposes because of the context, just reviewed, in which it occurs. Deutsch has important reasons other than his version of realism about quantum physics for thinking that science both properly inspires and can be unified—to scientific, not merely philosophical, purposes—by a naturalistic metaphysic that closely resembles that of Ladyman and Ross (2007). He also shares our view that the development of such metaphysics is important to a full-throated defence of Enlightenment progress based on science against conceptual and related forms of conservatism. But his devotion to Popper blinds him to crucial insights from anti-realist strains in the history of science, and this in turn ensnares him in a surprising form of conservatism when he seeks to metaphysically lever what anti-realists would regard as his favoured interpretation of quantum theory—the Everett multiverse—above all its rivals. Deutsch does not see matters that way; as far as he is concerned, Everettianism is not an interpretation of quantum mechanics but is simply the true fundamental physical description of the world.

From the fact that science in general has metaphysical implications, it does not follow that every episode in the history of science is at least an implicit exercise in metaphysics. This does not seem to occur to Deutsch. In his book he ignores episodes of the kind emphasized by philosophical anti-realists in which, according to their accounts, progress was made by scientists seeking only to develop more empirically adequate models of phenomena. Many examples can be cited. Entropy, as originally introduced, has been interpreted as a purely formal device without a physical interpretation. Heisenberg's matrix mechanics was explicitly instrumentalist in its creator's conception. Modern economists do not typically imagine that anything in a person's psychology corresponds to a utility function, and philosophers of economics who nevertheless appreciate the value of the construction therefore often offer a formalist or instrumentalist account (an example is Binmore, 2009). Contemporary physics introduces all manner of quasi-particles and other entities within models that anti-realists—including both philosophers and some physicists—understand as convenient fictions. Anti-realists might also object that Deutsch neglects the fact that ontological interpretations of theories and theoretical elements change over time, and theories are readily usable by later scientists who disregard the original ontological interpretations of them So, for example, the Clausius cycle in thermodynamics was originally conceived

in terms of the flow of material caloric particles, but the equations for heat flow are still used today without any such interpretation. Newtonian mechanics is now usually interpreted instrumentally—Deutsch says explicitly that there is no force of gravity (p. 107)—but it is highly predictively accurate and no less counts as science for that. It is important to note too that Newtonian mechanics had plenty of Deutsch's celebrated 'reach'. It is our first great example of unification in science, and it produced predictions that are accurate to one part in 10^6, from an evidential basis that was accurate to only one part in 10^3. Hence, it passes Deutsch's test for realism but its ontology as understood in a standard realist construal cannot be taken as literally sound.

For reasons to be explained below, we think that the standard terms of the realism-antirealism debate distort the history of science and, more specifically, its relevance to metaphysics. But let us defer our more radical critical perspective for the moment. Even someone who accepts the terms of the debate might be led by careful historical reflection to deny that either realists or anti-realists can reasonably be expected to offer the most compelling account of each and every episode of theoretical interpretation in science. According to this hypothetical voice of moderation, we do better to think of multiple specific debates as arising across the range of sciences, with realists winning some skirmishes and losing others. Sometimes a fight seems settled for one side for some time, and then the rival interpretation unexpectedly enjoys a renaissance. Sometimes there is no argument in the first place. All naturalistic philosophers, and philosophical scientists like Deutsch, should take these debates seriously because they arise and are resolved within science.

Some realist-antirealist arguments are of course more historically important than others. The great episode in the twentieth century was the long argument between Einstein, on the one hand, and Bohr and others on the other hand, over the measurement problem and the reality of the 'collapse' of the wave function in quantum mechanics (Fine, 1986). Deutsch sides firmly with Einstein metaphysically—not in the sense of denying the no hidden variables theorem, but in the sense of holding that unreduced stochastic generalizations indicate incomplete explanation. This leads him to defend the Everettian multiverse as established (though fallible) physics, partly on grounds that, among accounts that avoid hidden variables, it alone allows us to account for the observed statistical patterns in measurements taken within the frame of a single history, while maintaining that causal determinism holds across the multiverse. These

philosophical considerations are not the only basis for pursuing Everettian interpretations of quantum measurement; there are also important considerations that would generally be regarded as more narrowly physical and non-philosophical (Saunders et al. 2010; Wallace 2012). We will return to this point later. For the moment, however, let us summarize the logic of Deutsch's explicit philosophical argument as given in *The Beginning of Infinity*:[6]

(i) Either we must be instrumentalists about quantum mechanics or Everett is right.
(ii) Instrumentalism rests on sharply distinguishing between theories as formal devices and physical interpretations.
(iii) The distinction between theories as formal devices and physical interpretations has no sound basis in scientific practice.
(iv) Therefore we should not be instrumentalists about quantum mechanics.
(v) Therefore Everett is right.

Among the various ways of resisting this argument, one could accept Deutsch's premise (i) and deny premise (iii), thus blocking the inference to (iv) and (v). We expect that this would be a stance favoured by some philosophers and theoretical physicists. Experimentalists would be less attracted to it. In the next section, we argue against premise (i). We also supplement the argument in the previous section that complicates premise (ii). Premise (ii) is correct only given an assumption we deny, namely that the logical foil of instrumentalism is realism about the independent existence of types of objects quantified over by theories. For these reasons, we conclude that Deutsch is wrong to think that rejection of crass instrumentalism entails affirmation of the Everettian interpretation. (We think that premise (iii) is correct and we return to this below.)

Deutsch is right that outright instrumentalism, of the kind that interprets scientific theories as purely 'as if' descriptions that facilitate practical control, is to be rejected. The overriding objective of science is to

[6] Wallace (personal correspondence) questions the seriousness with which we consider Deutsch's arguments advanced in a book intended for a popular audience, in contrast with his work in the professional physics literature. However, it seems to us natural and appropriate that Deutsch should address his overtly metaphysical ideas to a broader readership; and our topic here are those ideas. Deutsch nowhere suggests that his philosophical arguments should be regarded as casual weekend work that should be taken less than fully seriously.

accurately model objectively real structures. Further purposes for such modeling, including explanation and out-of-sample prediction, are pursued with varying degrees of emphasis among specific scientific projects. But across-the-board instrumentalism should not be conflated with the sensible understanding that scientists often build theories, to excellent purpose, that isolate important aspects of real structure by resort to pretense that other aspects do not exist or do not matter.

The contentious parts of standard scientific realism are the semantic and epistemic doctrines that have been associated with it. The former holds that accepted theories should be taken as 'literally referring'[7] to unobservable objects. There is an important sense in which we reject this claim. Ladyman and Ross (2007) argue that what science aims and often succeeds at truly characterizing are structures rather than individual objects, observable or otherwise. To the question 'Are there really protons?' the answer is 'Yes', because theories in which protons are elements characterize real structure; but it does not follow from this that the world is partly composed out of individual pieces that intrinsically bear properties corresponding to predicates of the word 'proton' as it occurs in natural-language paraphrases of theories in mathematical physics. The epistemic aspect of standard scientific realism consists in the claim that we can and do have knowledge of unobservable aspects of the world (in addition to knowledge of observable aspects). If this is interpreted in terms of unobservable kinds of objects then it is dubious for all the reasons we have argued. However, understood in terms of claims such as 'The Kreb cycle is an unobservable structure that underlies cellular metabolism', or 'Mobile telephones work because microwaves propagate between them and telecommunication masts and satellites', where 'microwaves' describes recurrent structure for propagating a certain class of influences, then we do have knowledge of the unobservable.

There are other important components to realist thinking that we also endorse. These are as follows:

- The novel predictive power of scientific theories and the degree to which they unify the phenomena and apply to domains beyond those

[7] The phrase 'literally referring' involves a redundancy unless one thinks there is such a thing as metaphorical reference. Ross (1993) argues that there is such a thing, but that metaphorical reference operates at the level of whole structures rather than individual entities.

for which they were designed (reach) is surprising, and undermines across-the-board instrumentalism and motivates realism.

- The content of scientific theories must be understood modally, in so far as we are often able to advance scientific predictions about what would happen were distributions of specific measurement values to be somewhat different. What people call 'laws of nature' and 'causal laws', as well as generic and singular causal claims in science, correctly represent aspects of the structure of the world that go beyond regularities in the actual phenomena (see Berenstain and Ladyman, forthcoming).
- The distinction between the observable and the unobservable, as it figures in debates between advocates of standard scientific realism and advocates of standard instrumentalism, is of no general epistemological or ontological significance though it may sometimes be important within a particular science. Rejecting the significance of the distinction is an important expression of the rejection of anthropocentrism that Deutsch and we recommend. Note that the paradigmatic examples of unobservables are putative entities that are too small to be observed such as atoms, quarks, electrons, and so on. However, there are also examples of entities which are unobservable because they are too large. Consider tectonic plates that cannot possibly be surveyed and grasped in their full spatial extent in any single observation. More important are the domains of science, such as economics, evolutionary biology, or the behavioural sciences in which the standard distinction between the observable and the unobservable makes little sense at all. Utility functions, mating strategies, fitness, and deep grammar are unobservable only in the sense that to suggest that they are observable is a category mistake akin to wondering whether time or numbers can be observed. Chairs, people, tectonic plates, protons, and mating strategies are all instances of real patterns; and that is all there is to the objective existence of anything.[8]

[8] The force of 'objective' here is to exclude existence as aspectual of what Ladyman and Ross (2007) call 'notional worlds'. Spiny Norman exists in a second-order notional world constructed by Monty Python's Dinsdale Piranha, in which some properties ('is a hedgehog') apply to him and others ('is a penguin') don't; similarly, Dinsdale himself is a first-order real pattern in a notional world constructed by the Pythons, where he too is subject to accurate ('nails people's heads to tables') and inaccurate ('is a Welshman') characterizations.

- Science is progressive and cumulative in respect of more than its predictions about phenomena. In particular, the structures identified by earlier successful theories are preserved at least in approximate form in future theories (the general correspondence principle). Such claims if interpreted as being about individual kinds of objects are unlikely to be preserved without substantial modification if not wholesale abandonment across intertemporal theory adjustments. Theories and models also undergo revisions in the scopes of their sound application. As Deutsch says about Newtonian mechanics (p. 112), the inverse square law of gravitation is approximately recovered as a low energy limit in General Relativity. However, the predictions of Newtonian mechanics about the phenomena in regimes of the very fast, or the very small, are not even approximately preserved, for quantum tunnelling, time dilation and the equivalence of mass and energy are flatly inconsistent with Newtonian mechanics. Hence the cut between what is preserved and what is not is not the same cut as between the parts of the theory that describe phenomena and those that describe inferred structural relationships and limits.

Very importantly, the distinction between what is preserved and not preserved in theory change also does not match any distinction between the formalism and interpretation of theories. This is only to be expected in light of the correctness of Deutsch's third premise in the argument, viz.

(iii) The distinction between theories as formal devices and physical interpretations has no sound basis in scientific practice.

The best series of demonstrations of the soundness of this premise that we know of is given by Wilson (2006). However, note that an instrumentalist about Newtonian mechanics can readily admit that to use it requires interpreting the mathematics in terms of physical concepts such as mass, energy, and force. Hence, Deutsch's premise (ii)

(ii) Instrumentalism rests on sharply distinguishing between theories as formal devices and physical interpretations.

is wrong if intended to apply to specific theories. Intended as a claim about across-the-board instrumentalism, it is less easy to directly refute by citation of examples—it is hard to see how a general instrumentalist could *avoid* believing in pragmatically useful but ontologically empty formal

theories. But it is not clear that across-the-board instrumentalism makes coherent sense except as a generalization of claims about specific theories.

Insofar as instrumentalists deny the realist claims above we have already made clear where we disagree with them. However, we have also made clear that we think they are right to be sceptical about the literal truth of claims made about the natures of unobservable objects. That is to say, they are right to reject standard scientific realism as a basis for populating ontology, insofar as that ontology consists of theory-independent objects rather than Real Patterns. The idea of a 'theory independent' Real Pattern makes no clear sense unless theories are understood purely as representations. However, instrumentalists rightly stress that theories are also tools for intervention (Hacking, 1983).

Instrumentalists are also justified in their scepticism about merely explanatory theories, in emphasizing the discipline of subjecting theories to empirical testing, and in insisting that speculations about mechanisms must be continually reigned in by it. The distinction between what can and cannot be measured is epistemologically significant, and instrumentalists are right to emphasize that scientific theories must make contact with measurements to avoid the status of mere speculation. It is right to reject the demand for further explanations at certain points in the development of science, especially those that proceed by postulation (Ladyman and Ross, 2007: 61, cf. van Fraassen 2002: 37). For example, we deny the need to answer questions like 'What is a utility function?' if they are construed as requiring answers beyond descriptions of what utility functions look like mathematically and of how they figure in economic explanation and prediction. We return to this case below. The same point can be made about Bohr's version of the Copenhagen interpretation of the wave function; so 'realists' about the wave function are not entitled to dismiss Bohr's interpretation on general philosophical grounds. This is another example about which we will shortly say more. Furthermore, instrumentalism fits naturally with falliblism and the idea that we may be effectively infinitely far from knowing the full theoretical truth about the world. This is in accord with Deutsch's idea that we stand at the 'beginning of infinity'. In the words of James Jeans:

> Physics and philosophy are at most a few thousand years old, but probably have lives of thousands of millions of years stretching away in front of them. They are only just beginning to get under way, and we are still, in Newton's words, like children playing with pebbles on the seashore, while the great ocean of truth rolls, unexplored, beyond our reach (1943: 217).

Instrumentalism is important for deflating epistemic hubris and reminds us that, in the case of metaphysical questions concerning the ultimate form and nature of space and time, determinism, the origin of the universe, and so on, the jury is still out. Scientists who rush to pronounce on such questions in the light of the latest theories go beyond the evidence. Instrumentalism provides an important counterpoint to the claims of string theorists certain that there are eleven dimensions, or a cosmic landscape with multiple universes with different values of the fundamental constants, and so on; for all we know they will meet the fate of Newton's absolute space, Lavoisier's caloric, and Maxwell's ether.

Finally, we observe that just as Deutsch is wrong to dismiss instrumentalism as altogether without warrant or insight, he is also wrong to dismiss the idea that scientific knowledge is sometimes based on the mere generalization of an observed regularity. While it is certainly true that not all science is like this, there are cases that do conform to the simple instrumentalist view of what theories are. For example, it was surely known that Aspirin cures headaches before anyone had any account of the mechanism by which they do so or indeed any account of what Aspirin's molecular structure is. Similarly, much knowledge in science is generated purely by pattern recognition or regression analysis of large amounts of data, and indeed this is increasingly the case. Whether or not it is always better to have explanations, as Deutsch stridently and repeatedly insists, it is not the case that we have no knowledge without them.

4. Quantum metaphysics

There are two notions of ontology in play in discussions about scientific realism. The first concerns which entities scientists should be construed as 'really' believing in. Scientists regularly engage in such debates, wondering whether routine quantification should include some disputed class of entities. (For example, addiction theorists, having achieved a loose consensus that gambling is potentially addictive, are currently contesting the question of whether 'there is such a thing as' sex addiction; see Ross et al. 2008.) This sense in which science populates ontologies is drained of metaphysical significance if one maintains, following Ladyman and Ross (2007) that classification of entities into sets is mainly an exercise in practical book-keeping *except* insofar as it is driven by mathematicized theoretical modelling of structural patterns, in which entity language is typically

optional in principle. The second notion of ontology concerns the fundamental structure of the world. Deutsch conflates realism about scientific ontology in the first sense with realism about it in the second. Ontology in this second sense—which is equivalent to naturalized metaphysics according to Ladyman and Ross (2007)—necessarily involves one in reflections on quantum theory, because no other currently mature part of science is reasonably intended to restrict all possible measurement values in the universe at all scales.

There are two fundamental and intimately related problems in the philosophy of quantum mechanics. One is to say something sensible about systems in superpositions with respect to observables like being on one side of a box or another. The problem asks for an interpretation in the sense of an account of what the world is like when the quantum mechanical wave function gives it a bimodal probability amplitude over two disjoint regions of space. Wave functions are poised between physics and mathematics. Since the beginning of quantum mechanics they have been regarded on the one hand as fictional mathematical entities that represent probability amplitudes, and on the other as physical objects that can propagate, interfere with each other and in certain circumstances be partially localized as wave packets, or even so localized as to appear as an event at a point, or more generally to 'collapse'. This brings us to the second problem, the measurement problem: how do superpositions turn into the kind of relatively definite states we observe? Fuzzy and unsharp states do not help with dichotomous cases like cats being alive or dead.

It is often assumed that realism about the theory requires both an ontological interpretation of the wave function and perhaps some distinct particles or matter of some kind, and a solution to the measurement problem provided in part by that ontological interpretation. Everettian quantum mechanics is essentially a strong form of realism about wave functions, but not about their collapse. Bohm theory also requires realism about wave functions, albeit wave functions not construed as inhabiting space-time in which particles follow their trajectories. The Ghiridi, Rimini, and Weber (GRW) interpretation has spontaneous collapse and has been given an explicit ontological interpretation (Allori et al. 2008). Purely epistemic interpretations have no ontology because wave functions are construed as embodying our knowledge in some way, and so there is

no problem about them collapsing when we find something out in a measurement.[9]

Deutsch argues that realism about quantum mechanics means realism about the wave function, and then infers many worlds from that, with the wave function representing distinct branches each of which can be understood classically. The idea that the description the wave function gives is a description of an aggregate is the price that must be paid for maintaining a classical picture of individual branches and hence of the fundamental ontology. However, it is possible to be a realist in many ways without being committed to the wave function except as an abstract characterization of some—arguably fundamental—structural parameters that apply to the universe in general. Again we must distinguish between standard scientific realism and structural realism of the kind to which we subscribe. The standard realist might apply her realistic attitude to the apparatus of the particle accelerator, and to the classical fields that one uses to manipulate quantum systems. She may be a realist in the sense of believing that the energy of quantum systems is physically real, and that their states represent quantities like position and momentum. Indeed, since position and momentum can be weakly measured and both known reasonably precisely simultaneously without violation of the uncertainty relation, she may be a realist about the average and approximate values of those physical quantities and associated physical processes such as motion or tunnelling or diffraction. She may believe that there really are emergent particles such as protons, photons, and electrons and that she is able to manipulate them and interact with them in the laboratory. She may believe that she can manipulate wave packets. All of this is consistent with her taking a standard instrumentalist attitude to the wave function and regarding it as non-referring formalism. This is the position against which Deutsch aims to persuade her.

Recall that Deutsch's defence of Everettian realism about quantum mechanics depends on his insistence that we cannot make sense of the

[9] See the essays in Ney and Albert (eds.) (2012) for discussion of wave function realism from a variety of perspectives. Recent work in the foundations of physics (Pusey et al., forthcoming) suggests that a purely epistemic view of the wave function is untenable if we are to be realists at all. The conclusion of this paper is roughly that if there are hidden variables then the wave function itself must be among them. The proof seems to involve a separability assumption that could be contested, and one could of course also contest the realism that is presupposed, but in any case realism about the wave function of the kind established does not imply the kind of realism we find objectionable in Deutsch's defence of Everett.

theory in instrumentalist terms as a bare formalism. Albert (1996) goes so far as to claim that in the absence of a definition of measurement from which one can deduce when collapse occurs, quantum mechanics has no empirical consequences. He can only intend this as some philosophical version of 'strictly speaking', since clearly physicists have been drawing empirical predictions from quantum mechanics for years very successfully despite the lack of a viable solution to the measurement problem. One need not take them to be instrumentalists in Deutsch's sense to make sense of this. The formalism of quantum mechanics is replete with physical content. To begin with we have Planck's constant everywhere telling us that we are dealing with the quantum of action, where the latter is a familiar physical component of classical mechanics. Similarly we have Hamiltonians, linear and angular momentum operators and spatial co-ordinates, all of which have physical content. There is a wealth of physical content distributed throughout quantum mechanics in the absence of an account of measurement. For example, atomic chemistry and the structure of the periodic table are recoverable from the quantum description of stationary states of electrons and the structure of angular momentum eigenvalues, independently of any story about collapse of the wave function.

This is roughly the sort of account favoured by earlier versions of the Copenhagen interpretation. Later versions, under the influence of Pauli and von Neumann, *did* include a story about collapse, but interpreted it as a consequence of measurement. This was arguably an abandonment not so much of realism as of naturalism itself, and produced as a backlash the near-consensus among philosophers of physics that the only viable options on the table are those (e.g. Everett, Bohm, GRW) that 'explain' collapse in Deutsch's sense. But Bohr's version of Copenhagen denies precisely this need. In the statistical distributions of quantum states, some are wave-like and some are particle-like, and we can say a great deal—to a remarkably precise extent—about the real patterns that predict these distributions both in and out of samples. Statistically, the particle-like states are special cases of the wave-like states, but the former admit of no general mechanical or dynamical explanation in terms of processes that transform the latter.

This does not mean that the wave function is 'pure formalism' in the Bohr interpretation. If one eliminated the physical content from quantum physics and ended up with only the mathematics of Hilbert spaces and statistical operators, one would be far short of Copenhagen or any other quantum physics—trivially, for one would have removed the physics.

Though there is a purely mathematical residue of quantum physics, it is not clear that there is a physical residue when we eliminate the mathematics from quantum physics. There are very broad qualitative claims such as the quantization of energy, the superposition principle and the uncertainty relations, but expressing them requires mathematics. The distinction between Everettian, Bohmian, and other interpretations of quantum mechanics and uninterpreted formalism is a false dichotomy. Denying the intelligibility of a bare formalism and interpreting quantum mechanics in physical terms does not imply that one must have a solution to the measurement problem, whether Everett's or an alternative.

It is instructive to note that Deutsch does not discuss quantum field theory (QFT). The present state of the latter makes denial of standard realism much more plausible than Deutsch allows. In general, quantum field theories are only able to describe the effects of the forces between particles by the imposition of length-scale cutoffs to the contributions to scattering amplitudes due to very short and/or very long-range interactions. Quantum field theories form a hierarchy where each is an approximation to the theory below without any fundamental level to provide the foundation. Of course, some such theory may yet be found, but all the existing theories are so-called 'effective theories' that include approximate degrees of freedom in the form of 'particles' that emerge from a substructure that is ignored (Wallace, 2006).[10] A structural realist view, at odds with both standard realism and standard instrumentalism, about the ontology of QFT seems completely naturalistically appropriate, and does not amount to regarding QFT as a bare formalism. (More generally the fact that physical theories are usually applicable only within certain regimes, and the emergence of effective degrees of freedom in the form of particles, fits very well with Ladyman's and Ross's [2007] thesis of the scale relativity of ontology.) Furthermore, QFT and the standard model incorporate symmetry groups and use them to arrive at empirical predictions for particle families that are completely independent of a solution to the measurement problem.

Everett was dissatisfied with the Copenhagen theory of measurement, particularly, with its ambiguity as between the Bohrian refusal to discuss

[10] Wallace as an Everettian would of course not endorse the use to which we are putting his account of quantum field theory. He regards Everettianism as having the advantage that by not being revisionary physics it fits well with quantum field theory, unlike Bohm theory and dynamical collapse theories.

particles as if they were things in themselves, and the seemingly realistic construal of wave functions inside systems collapsing as if the latter was a physical process. Ghiridi, Rimini, and Weber showed that such models can be constructed and there have even been relativistic physical collapse models. Bohr's most basic reaction to quantum mechanics was to abandon some aspects of the strong form of realism upon which Einstein insisted. We now know, because of the experimental violation of the Bell inequalities, that locality and the reality principle that measurements reveal antecedently and independently possessed values cannot both be true. However, as d'Espagnat (2006) points out, nobody has proved a version of Bell's theorem using only locality and not some form of realism, or counterfactual definiteness about unperformed measurement results. D'Espagnat himself is sympathetic to a Bohrian or neo-Kantian version of structural realism. He argues that quantum physics undermines not metaphysical realism, but any confidence that we are able to grasp the ultimate 'ground of things' (see especially 2006, Chapter 19); we would strengthen this by denying that there is a convincing basis for believing there is any such ground.

We do not deny that Everettian realism is an option to be taken seriously and further developed. It is clearly motivated by considerations arising within physics, and offers the potential to unify some otherwise disconnected parts of physics (Wallace, 2012). Our point is only that one need not insist on a realist picture of what happens in quantum measurements in order to avoid instrumentalism about quantum physics more generally. D'Espagnat (2006) insists that decoherence theory goes some way towards explaining how the classical and relatively determinate macroscopic world emerges from the quantum realm despite the fact that we cannot interpret it literally as a causal account of how wave functions collapse. Indeed, he insists that quantum physics represents the world as if it were separable into distinct systems where entanglement tells us that ultimately it is not.

For many who are convinced of the indispensability of picturing distinct individuals separable in space as a basis for scientific metaphysics, it seems more acceptable to abandon locality. Albert and Galchen (2009) argue that locality is in any case incompatible with quantum mechanics. However, there is no overt conflict between quantum mechanics and Special Relativity and furthermore, there are empirically successful relativistic quantum mechanical equations such as the Klein-Gordon and Dirac equations. The

pursuit of Lorentz covariance of fundamental equations such as the Lagrangians for different interactions is a major methodological pillar of QFT and its astonishingly accurate predictive success. The Dirac equation, for example, is needed to explain why gold is gold and not silver as the Schrödinger equation would have it. Again the everyday application of these equations to condensed matter physics does not rely upon a solution to the measurement problem.

In light especially of entanglement and its implications, we view Everettian realism as a better motivated unifying theory than Bohm theory, which is based partly on insistence on classical trajectories in space and time, by determinism and by the insistence that physics be interpretable in terms of individual, separable objects. However, the Everett interpretation, while not revisionary physics, is (naturalistic) *metaphysics*, in that it unifies theories indispensible for predictive success, but contributes no novel predictions itself. Our point is only that an alternative unifying metaphysics is available. This involves denying that there is any real causal process corresponding to wave function collapse, and hypothesizing that the statistics directly identify the relevant real pattern—that is to say, a basis for prediction that supports all relevant counterfactuals and is non-redundant in the specific sense that no 'deeper' pattern—Everettian or otherwise—supplants it.

Taken more or less at face value, our best fundamental physics tells us that there are no little things. Consider again QFT, in which particles are excitation modes of fields, that themselves are assignments of operator-valued measures to regions of space-time. Particle number is frame dependent, and every generation of particles turns out to be a collection of effective degrees of freedom that approximate the structure of the underlying deeper field theory. Throughout field theory and condensed matter physics there are quasi-particles—things that look and behave like particles but may be aspects of sound waves, or propagating oscillations of some kind, and from the perspective of each more fundamental theory, allegedly fundamental particles are quasi-particles. In ordinary non-relativistic quantum mechanics the consensus is that particles lack the attributes to qualify as what is classically meant by individuals since they fail to have properties sufficient to absolutely discern them from each other. The existence of entangled states, as well as their lack of determinate spatiotemporal trajectories further shows that thinking of them as little things is entirely inappropriate. When it comes to space-time physics we

have learned that the identity and individuality of space-time points is grounded in the metrical relations between them and not the other way around. The situation in other sciences is often similar. Biological individuality is relative to selection in the sense that to determine what biological individuals there are in some domain it is necessary to see what counts as an individual for the purposes of Price's equation. So cells are not biological individuals in multicellular organisms but they are when they act alone. The message in general is that the idea of distinct individuals is a pragmatically justifiable one for the purposes of scientific representation and referential bookkeeping in practical contexts, but it breaks down when we are rigorous about restricting attention to the non-redundant patterns with greatest reach (to borrow Deutsch's term just as he intends it). Deutsch's Everett branches are each construed as classical worlds with particles following trajectories. While acknowledging some unifying value delivered by this specific interpretation of the wave function, the sacrifice it implies of other phenomena—particularly quasi-particles and entanglement—from the unified picture is reason to decline to endorse it.

According to our notion of fundamental physics what is fundamental about it is not its ontology in the sense of which kinds of entities or referents it quantifies over. It is important to ontology in an alternative sense of showing which constraints stand on the boundary of naturalized metaphysics by imposing consistency restrictions on all hypotheses in all sciences.

Wave functions as devices for describing structure are not unique in this character; we find similar devices throughout the sciences. Consider utility functions as used in economics. In their most methodologically robust deployment they constrain the structure of estimations of patterns in choices distributed across populations of agents. A typical economist is a realist about choices and people. In wondering about the effect that a change in incentive structures will have on the statistical patterns in choices distributed over a population of people she will need to estimate distributions of risk preferences—that is, levels of aversion and attraction to risk itself—and discount rates applied to rewards received at different temporal distances from choice points. She will achieve accuracy in her predictions only if she jointly estimates the risk preferences and discount functions, using a model that allows for heterogeneity in the population, not only with respect to numerical values but with respect to functional forms (Andersen et al. 2008). None of the individual coefficients that enter

into the computation of the pooled estimated output should be imagined to be elements of any person's psychology, or even to be standing properties of particular people; they are summaries of statistics produced by earlier stages of data analysis that enter into later stages. A crucial set of parameters that must be calculated between risk preference estimation and joint risk-and-time preference estimation is the shape (that is, the curvature) of each member of the set of heterogeneous utility functions. It would be absurd to be a 'realist', in the standard philosophical sense, about this curvature and wonder which independent objects are so curved. Yet clearly the utility function describes real structure in the statistical dispositions of the population to respond to the incentive shift; that, after all, is the point of the exercise, and there is certainly clear meaning to the possibility of *mis*estimating utility function curvature, which might be the reason why some predictions come out wrong. Note that, as with wave functions, estimation of individual utility functions is not wrong-headed a priori, and indeed economists often do this. But the stability of individual utility function estimations from one measurement episode to the next is typically much attenuated over time compared with estimations based on pooled data.

In the framework of dichotomous opposition between standard realism and standard instrumentalism, wave functions and utility functions—along with countless multitudes of similar devices used across the sciences—are liable to be regarded as 'pure formalism'. But this fails to do justice to the ways in which they are embedded in theoretical and experimental contexts that are rich in ontologically robust structure. Even in the most abstract reaches of formal science, where the only objects of manipulation are representations of functions or elements of topology or high-dimensional sets, there is always ontological commitment to the data and their statistical structures.

5. The world is the totality of non-redundant statistics

Deutsch joins distinguished company including Einstein and Kant in taking for granted that any irreducible stochasticity attributed to a class of systems constitutes an undischarged call for explanation. If this is granted, then to assert that the call *cannot* be discharged is to deny Enlightenment optimism, by insisting that there is at least one insoluble problem.

The principle that reality cannot be irreducibly stochastic because to allow that it might be is methodologically unacceptable, is *not* consistent with a naturalistic attitude to metaphysics (or epistemology). It is, instead, an instance of dogmatic metaphysics—to borrow a phrase of Kant's more or less exactly as Kant intended it. This lapse is so sharply inconsistent with Deutsch's otherwise rigorous commitment to drawing his premises from contemporary science that it itself calls for an explanation, in the sense of a diagnosis. This takes us directly into a core question at the heart of the project of naturalizing metaphysics, since it would isolate a fundamental ambition of non-naturalistic metaphysics that a naturalized metaphysics should discard.

A key problem in Deutsch's particular approach to the problem arises from his casual understanding of explanation, which invites ambiguity. The most common interpretation of explanation in both everyday semantics and the philosophy of science is that an explanation relieves some element of surprise by attention to a previously undisclosed fact that, in conjunction with other things known, turns the surprising state of affairs into something that might have been expected after all. An explanation in this sense is the provision of a missing premise in an implicit argument. We will suggest that this isn't generally what Deutsch means by explanation; but then his case against irreducibly stochastic laws of nature[11] rests on equivocally appealing to this standard idea of explanation after all.

We do not deny that scientists often give explanations in the standard sense. For example, if someone cooled some water to a temperature well below zero degrees Celsius and, due to osmotic depression, encountered a failure to freeze that baffled them, then a scientist could relieve their wonder by supplying relevant details about the physics of state equilibrium transitions and evidence of a high proportion of dissolved ions in the sample in question. Note that this unremarkable kind of example asks us to imagine the interaction of a scientist with a layperson or student, or of a scientist who knew the history of the sample with a scientist who didn't. Cases of this sort are not templates of the way in which scientists generally make new discoveries. This is indeed a point that Deutsch makes repeatedly in his book. Popper's best motivated complaint against common-and-garden

[11] On some conceptions of 'laws of nature', the idea of an irreducibly stochastic law is self-contradictory. As an analytic dogma, however, this can have no persuasive force for a naturalist. For our naturalistic interpretation of laws, see Ladyman and Ross (2007: 281–90).

versions of verificationism is that it is a distortion to imagine scientists as generally going about with settled pictures of reality and seeking to fill in missing premises in argument patterns so they can apply these settled pictures to newly observed bits of the world. That is rather more like the way in which philosophers do their business.

But what is the more accurate story? A problem with Deutsch's over-reliance on Popper is that one finds nothing in Popper's philosophy that transparently points to a clear alternative meaning for explanation. We can, however, find one in the work of a much more sophisticated philosopher of science whom Deutsch never mentions: C. S. Peirce.

We will work up to the value that Peirce adds here by retaining Popper as the critical starting target. Popper's historical context led him to routinely contrast his falsificationism with the verificationism of the positivists. The falsehood of a generalization, he famously emphasized, could be deductively inferred from a single contradictory observation, but cannot be 'confirmed'—that is, in Popper's (and Deutsch's) tendentious understanding, conclusively shown to be true—by any finite number of observations. Critics (e.g. Hempel, 1945) quickly noticed, however, that this asymmetry between falsification and confirmation is parasitic on an entrenched semantic habit, that of stating generalizations as affirmative associations of objects and attributes rather than negative ones. If people had the opposite habit, then the logical empiricists could have turned Popper's argument around and used it against him. Now, this familiar criticism is not the point of interest here. The point, rather, is that Deutsch inherits from his philosophical mentor the tendency to see the whole epistemology of science as a war between ideologies of 'deductivism' and 'inductivism'—views that promote deduction or induction, respectively, as the methodological foundations of scientific reasoning.

Deutsch does not rest his main critique of induction on the simple mistake that, as we have pointed out, he shares with Popper of thinking that justification must aim at certainty. Instead, he argues as follows. A useful inductive argument seeks to associate a degree of probability with a generalization. If by probability we mean to refer to subjective confidence, however, then inductivism makes scientific knowledge[12] relative to the psychological histories of particular reasoners, and directly

[12] By 'scientific knowledge' we allude only to the network of established scientific belief that is treated for practical purposes as settled. We do not refer to the analytic philosopher's

undermines its objectivity. If, more hopefully, we understand probability as relative observed frequency in draws from a sample, then statistical theory tells us that induction understood in these terms is a tool of sharply limited reach. This is because, in any instance of application, confidence in avoiding estimation bias requires that we understand the mechanism that generated the sample from which our observations are drawn. This in turn suggests that induction only soundly applies to phenomena that we already substantively understand, in the sense of having a good model of their structural parametrization. Thus induction, Deutsch concludes, cannot be the basis for the ever-widening reach of scientific theory.

This criticism of induction as a plausible basis for *profound* scientific learning—it being no critique of induction as a means to learning variable distributions in well-modelled domains—was anticipated by Peirce. We draw here, and in the comments on Peirce to follow, on Hacking (1990, Chapter 23). Peirce, Hacking explains, saw that induction serves the amplification of scientific knowledge not by attaching specific probabilities to specific conclusions, but by reassuring us that methods of investigation that have worked more often than not in the past are good guides to belief in the present—so long, that is, that we understand what those methods actually were. So far as inter-theoretic comparison is concerned, induction is then best understood as a fundamental *practice* supporting second-order confidence in theories with good track records, as opposed to first-order knowledge encapsulated in theories themselves.

Peirce had two names for the kind of procedure that, according to him, *does* yield qualitative amplification of knowledge. To emphasize that this is something apart from both induction and deduction he coined the term 'abduction'. His original and more descriptively effective term, also favoured by Hacking, was 'hypothesis'. Peircean hypotheses are in fact Popper's famous 'conjectures': generalizations worth further investigation because, if they are borne out, they structure ontologies of sample-generating processes for which we can then compute frequency distributions of variables we want to predict or control. Hypothesis, in Peirce's sense, is what Deutsch generally means—except when he slips in promoting dogmatism about the Everettian multiverse—by 'explanation'.

justified true beliefs that can never actually be identified as a specified subset of the justified beliefs.

Hacking himself reminds us how close everyone stands to the precipice over which Deutsch slips when he says—writing twenty-one years ago—that 'a few philosophers follow Gilbert Harman's attractive phrase "inference to the best explanation"' (1990: 207) as a third alternative to Peirce's two terms. Harman's phrase is harmless in its literal semantics. However, in the majority of philosophical treatments that rest major weight on IBE, 'best explanation' is typically read as 'most intuitively satisfying postulation'. As Ladyman and Ross (2007, Chapter 1) argue at length—and as Deutsch seems to agree, except where quantum mechanics is concerned—this is explanation in the anti-scientific sense that encourages what we call 'domestication' of the astonishing discoveries science routinely delivers by trying to accommodate them within comfortable conceptual frameworks. A prime example of domestication is conceiving of atoms as tiny, extended, distinct little 'things' that persist in space through time and 'give rise to' emergent macroscopic phenomena by changing proximity to one another and exchanging energy in the form of other, still littler, things. Quantum entanglement in particular and quantum physics in general, especially quantum field theory, show that there is *no sense at all* in which atoms or sub-atomic particles resemble little macroscopic things reduced drastically in size. In undomesticated physics, particles don't resemble any kind of entity that people had ever imagined prior to the twentieth century.[13] This is a decisive consideration in favour of anti-reductionism: there is no convincing reason to believe that the micro-scale mechanistic structure that the reductionist treats as explanatory bedrock exists.

The urge to domesticate is precisely the parochialism against which Deutsch wields his vigorous neo-Enlightenment rhetoric. As Ladyman and Ross (2007) show by reference to numerous instances, it consistently leads philosophers into the culturally disastrous game of urging outright resistance to new conceptualizations suggested by scientific discovery, in

[13] The 'bound state' of quarks that we call the proton has a mass of 938.27MeV even though the component quarks have a combined mass of about 15MeV. Hence, most of the mass of nucleons is due to the energy associated with the gluons that hold the quarks together and not the quarks that allegedly 'compose' the nucleons. In fact, quarks provide the structure of nucleons via their charge and other properties but not much of their substance. Quark interactions and how they behave at different length scales over time is what composition really amounts to in this case, hence, as Ladyman and Ross (2007) argue, composition is diachronic and scale-relative and due to the dynamics of interaction among the parts.

defence of the conservative's ever tedious 'common sense'. Analytic metaphysics applied to science is necessarily conservative in this way.

Deutsch's dogmatic refusal to allow for the possibility of irreducible stochasticity is similarly conservative. We are by no means the first philosophers to argue that refusal to allow irreducible stochasticity in the fundamental laws of nature is a domesticating move based ultimately on residual analysis of some general metaphysical notions, especially causation. Over a century ago, Peirce thought the same thing. He doubted deterministic foundations of reality for two main reasons, both based on fresh science of his day.

First, he was impressed by the power of evolutionary theory to explain permanent structural change. If such change is possible at the macro-scale, why should it not be possible at all scales? Earlier thinkers who had seriously entertained this question, such as Plato, responded by claiming that if nothing in reality is unchanging then knowledge is impossible. This objection, already highly questionable in resting a metaphysical principle on strong epistemological assumptions, collapses entirely once modern statistics arise, and shows us that we could have knowledge of the principles of frequencies and their estimation from data that holds good through any storm of fluctuation in contingent values. Ultimately, we can reconstruct causation in these terms, as asymmetric relations of statistically detectible influence among networks of variables. Hacking summarizes Peirce's view on the relationship between laws and fixed quantities as follows:

> Laws of nature are commonly presented as equations with some fixed parameters . . . But if laws are evolving from chance, we need not imagine that the constants are anything more than limiting values that will be reached in some indefinite future. The ultimate 'reality' of our measurements and what they measure has the form of the Gaussian law of error. (1990: 214)

Second, Peirce was interested—to the point of running his own experiments—in Fechner's psychophysics. This rested on the idea that judgements about measurable quantities are subconscious estimates, limited by predictable Gaussian error, that could in principle be performed by mechanical calculation so long as assumptions about distributions hold good. In Fechner's and Peirce's time the suggestion that brains could compute such estimates was pure speculation, the only hypothesis they could think of to explain experimental observations of people guessing the relative weights of objects. The hypothesis was testable, however, and now we can

say—using the language favoured by Deutsch if we like—that it gloriously survived efforts to falsify it. Using single-cell electrical recording, neuroscientists have observed groups of neurons in monkeys' brains computing near-optimal randomization of actions when rewarded for tracking patterns of light flashes varying in their stochastic properties through the experimenter's manipulation (Glimcher, 2003; Dorris and Glimcher, 2004; Lee et al., 2004; Lee et al., 2005). Furthermore, simple conditioned learning algorithms have been proposed which, supplemented by drift diffusion familiar from statistical modeling of diverse phenomena, would explain how clusters of such simple machines as neurons could compute these dynamic estimations (Lee and Wang, 2009). In a different extension, Fechnerian error is now a basic variable for estimation in economists' most sophisticated styles of modeling of human behaviour under risk and uncertainty (Harrison and Rutström, 2008).

The importance of Peirce's interest in psychophysics was, as Hacking explains, that it encouraged him to reconceive of properties of frequencies not as second-order properties of judgments, which people strive to bring into correspondence with fixed underlying constants, but as basic properties of the external world that constitute its structure. Remarkably, Peirce recognized in this idea the basis for a new conception of scientific method, the method that now overwhelmingly dominates the everyday life of the scientific community across almost all disciplines, of exploiting the systematic patterns of variation in large data sets generated by known processes to minimize estimation error—that is, to tell noise apart from structure. Where randomization with respect to a dependent variable in a model of such data can be experimentally imposed or instrumented, regression can even be used to discover causal relationships (Angrist et al., 1996; Angrist and Pischke, 2009)—on the Peircean understanding of causation indicated above. Such estimation is core activity in every science that relies on modeling quantitative data—including, in particular, quantum physics.

Deutsch, it will be recalled, vigorously defends anti-reductionism. However, his explanation for it, unlike Peirce's, is philosophical rather than empirical. Abstractions, he claims, have causal powers. We find this familiar term of philosopher's art unhelpful (see Ladyman and Ross, 2007: 4). Even so, we understand well enough what Deutsch is getting at: scientists often arrive at generalizations by, in part, imposing constraints derived from pure mathematics. (Consider, to take one of Deutsch's examples, the explanation for the fact that generations of cicadas appear

in cycles of years punctuated by prime numbers.) However, inferring a metaphysical principle such as 'no irreducible stochasticity' from an epistemological fact—that mathematics is efficacious in constraining scientific theories and models—is hardly in the spirit of realism.

We do not defend the Peircean alternative *because* we favour realism. That would be yet another way of putting the philosophical cart before the scientific horse. As discussed earlier, we regard the whole long realist-empiricist debate as highly instructive even as we endeavor to transcend it (Ladyman and Ross, 2007: 57–64). But the Peircean hypothesis we invoke to explain—in Deutsch's sense of 'explain'—the efficacy of statistical theory is the simplest one to which a realist should have recourse: the world is stochastically structured.

This hypothesis *explains* anti-reductionism but does not *motivate* it. Anti-reductionism is motivated, as an assumption regulating model construction, not conceptually or metaphysically but empirically. Reductions among domains of inquiry are extremely rare across the sciences, and in case after case we have strong explanations of why reduction is unachievable (Batterman, 2002). It follows from the truth of general anti-reductionism that, however completely the generalizations of fundamental physics constrain all measurements taken at all scales of real patternhood, one cannot hope to explain all or most real patterns by showing that they are determined by these generalizations. Were this *not* true, we would indeed expect that many putative 'higher level' patterns would be reduced away—eliminated—and thus turn out to be ontologically redundant. Most of these non-redundant non-reducible patterns are also irreducibly statistical, in the sense that they are generated by stochastic processes. We therefore should not expect that there should be explanations taking the form of systems of deterministic laws, for many of the phenomena we model most rigorously and reliably. Dissatisfaction of the Einstein-Kant type with irreducibly statistical fundamental physics, then, has no sound motivation in either epistemology or naturalistic metaphysics. The world is the totality of non-redundant statistics, not of things.

This is the key metaphysical conclusion to which empiricism as an epistemological program has always tended, even if most versions of empiricism have come up short of it because empiricists are loath to draw *any* metaphysical conclusions as a matter of (misplaced) principle. We would hesitate to regard the conclusion as vindicated were it not for the overwhelming empirical triumph of standard quantum mechanics.

Deutsch is right that empiricism does not tell the whole story. The realist's epistemological optimism about the boundless future of scientific progress may be expressed, in terms of the very empiricist principle just stated, as follows. We expect that there will always be new discoveries about convergences among all existing efforts to represent all the non-redundant statistics. Such discoveries will occur along the moving frontier of mathematical innovation that links models of specific classes of phenomena. At this frontier, the distinction between mathematics and empirical science is intrinsically unclear. Because this process amounts to the discovery of new fundamental features of reality, a structural realist, and a naturalistic metaphysician, should decline to try to give a *general* answer to the question as to whether the structure of reality is mathematical or physical.

It is sometimes useful to distinguish between purely formal and empirically interpreted parts of specific theories or models. When an economist, for example, constructs a so-called numeraire in which to represent the comparative values of all elements in a preference field, this should not be understood as reflecting any aspect of an empirical system being modelled. Examples of this sort are rarer than most philosophers have supposed, and as Wilson (2006 and this volume) demonstrates with a wealth of essential exemplary detail, the essential cunning tricks of applied mathematics typically make the boundaries between formalism and empirical description fuzzy. Yet even fuzzy distinctions have their uses. However, as one approaches the limit where the model in question is of the whole of reality—the domain of naturalized metaphysics—the distinction becomes ever more resistant to practical determination. Which parts of mathematics constrain all measurements of empirical magnitudes at all scales? One is inclined to hypothesize: less than all of mathematics, since mathematics is (arguably) infinite. But which parts might not be needed? There seems to be no basis for principled speculation, which is as good as saying that the value of the question collapses.

At this point it might seem that we have painted ourselves into a corner. This is indeed the impasse that some of our critics (e.g. Dorr, 2010) allege us to be stuck in. Structural realism, they say, in principle lacks resources for being able to distinguish pure formal structure from empirical reality, and turns into Pythagorean idealism. Mysticism is clearly not what most naturalists would regard as a happy final refuge.

However, there is a more plausible alternative that doesn't lead in this direction. The fundamental empirical structure of the world is not

mathematical but statistical. *And there is no such thing as purely formal statistics*. The 'principles' of statistics are simply whatever dynamics emerge from our collective exploration of, and discovery of patterns in, data. Applied scientists are continually reminded of this in their day-to-day work, especially as our new technological capacity at last realizes the potential for limitlessly evolving of co-adaptive growth of new techniques for discerning real patterns and new demand for statistical innovation arising from the discovery of new possible patterns. In light of this picture, Deutsch's Enlightenment optimism is maximally compelling. It is made *more*, not *less*, persuasive by the grandest discovery of the twentieth century, that fundamental physics, and therefore reality itself, are irreducibly statistical. What is the world? It is the endless weave of patterns to be extracted from noise, at an endless proliferation of mutually constraining scales, that we will go on uncovering forever as long as we have the collective and institutional courage that comes from love of objective knowledge, the great moral core of the Enlightenment.

References

Albert, D. (1996). Elementary quantum metaphysics. In J. Cushing, A. Fine, and S. Goldstein (eds.) *Bohmian Mechanics and Quantum Theory: An Appraisal*, pp. 277–84. Dordrecht: Kluwer.

—— and Galchen, R. (2009). A quantum threat, *Scientific American*, March 2009: 32–9.

Allori, V., Goldstein, S., Tumulka, R., and Zanghì, N. (2008). On the common structure of Bohmian mechanics and the Ghirardi–Rimini–Weber theory: Dedicated to Giancarlo Ghirardi on the occasion of his 70th birthday. *British Journal for the Philosophy of Science*, 59:353–89.

Andersen, S., Harrison, G., Lau, M., and Rutström, E. (2008). Eliciting risk and time preferences. *Econometrica*, 76: 583–619.

Angrist, J., Imbens, G., and Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91:444–72.

—— and Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton: Princeton University Press.

Batterman, R. (2002). *The Devil in the Details*. Oxford: Oxford University Press.

Bennett, C. (1990). How to define complexity in physics, and why. In W. Zurek (ed.) *Complexity, Entropy and the Physics of Information*, pp. 137–48. Boulder: Westview.

Berenstain, N., and Ladyman, J. (forthcoming). Ontic structural realism and modality. In E. Landry and D. Rickles (eds.) *Structural Realism*. Berlin: Springer.

Binmore, K. (2009). *Rational Decisions*. Princeton: Princeton University Press.
Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford: Oxford University Press.
—— (1989). *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.
—— (1999). *The Dappled World*. Cambridge: Cambridge University Press.
Dennett, D. (1991). Real patterns. *Journal of Philosophy*, 88: 27–51.
Deutsch, D. (2010). Apart from universes. In S. Saunders, J. Barrett, A. Kent, and D. Wallace (eds.) *Many Worlds: Everett, Quantum Theory, and Reality*. Oxford: Oxford University Press.
—— (2011). *The Beginning of Infinity*. London: Allen Lane.
d'Espagnat, B. (2006). *On Physics and Philosophy*. Princeton: Princeton University Press.
Dorr, C. (2010). Review of James Ladyman and Don Ross, *Every Thing Must Go: Metaphysics Naturalized*. *Notre Dame Philosophical Reviews,* 6 [online]. Available: <http://ndpr.nd.edu/news/24377/?id=19947> accessed 24 May 2012.
Dorris, M., and Glimcher, P. (2004). Activity in posterior parietal cortex is correlated with the relative selective desirability of action. *Neuron*, 44: 365–78.
Dupré, J. (1993). *The Disorder of Things*. Cambridge, MA: Harvard University Press.
Fine, A. (1986). *The Shaky Game*. Chicago: University of Chicago Press.
Friedman, M. (1999). *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.
—— (2001). *Dynamics of Reason*. Stanford: CSLI.
Glimcher, P. (2003). *Decisions, Uncertainty and the Brain*. Cambridge, MA: MIT Press.
Hacking, I. (1983). *Representing and Intervening*. Cambridge: Cambridge University Press.
—— (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.
Harrison, G. and Rutström, E. (2008). Risk aversion in the laboratory. In J. Cox and G. Harrison (eds.) *Risk Aversion in Experiments*, pp. 41–196. Bingley: Emerald.
Hawking, S. and Mlodinow, L. (2010). *The Grand Design*. New York: Bantam.
Hempel, C. (1945). Studies in the logic of confirmation. *Mind*, 54: 1–26 and 97–121.
Horgan, T. and Potrc, M. (2008). *Austere Realism*. Cambridge, MA: MIT Press.
Jeans, J. (1943). *Physics and Philosophy*. Republished 1981 by New York: Dover.
Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science*, 29: 409–24.
—— (2002). Science, metaphysics and structural realism. *Philosophica*, 67: 57–76.
—— and Ross, D. (2007). *Every Thing Must Go*. Oxford: Oxford University Press.

Ladyman, J. and Lambert, J., and Weisner, K. (forthcoming). What is a complex system? *European Journal* of *Philosophy of Science*.

Lee, D., Conroy, M., McGreevy, B., and Barraclough, D. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cognition and Brain Research*, 22: 45–8.

———————and —— (2005). Learning and decision making in monkeys during a rock-paper-scissors game. *Cognition and Brain Research*, 25: 416–30.

——and Wang, X.-J. (2009). Mechanisms for stochastic decision making in the primate frontal cortex: Single-neuron recording and circuit modeling. In P. Glimcher, C. Camerer, E. Fehr, and R. Poldrack (eds.) *Neuroeconomics: Decision Making and the Brain*, pp. 481–501. London: Elsevier.

Mäki, U. (1992). On the method of isolation in economics. *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 26: 319–54.

Ney, A. and Albert, D. (eds.) (2012). *The Wave Function: Essays in the Metaphysics of Quantum Mechanics*. Oxford: Oxford University Press.

Pusey, M., Barrett, J., and Rudolph, T. (2012). The quantum state cannot be interpreted statistically [online]. Available: <http://arXiv:1111.3328v1> [quant-ph] accessed 15 May 2012.

Quine, W. V. O. (1969). *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Ross, D. (1993). *Metaphor, Meaning and Cognition*. New York: Peter Lang.

——Sharp, C., Vuchinich, R., and Spurrett, D. (2008). *Midbrain Mutiny: The Behavioral Economics and Neuroeconomics of Disordered Gambling (Economic Theory and Cognitive Science, Volume Two)*. Cambridge, MA: MIT Press.

Saunders, S., Barrett, J., Kent, A., and Wallace, D. (eds.) (2010). *Many Worlds: Everett, Quantum Theory, and Reality*. Oxford: Oxford University Press.

Smolin, L. (2007). *The Trouble With Physics*. New York: Mariner.

van Fraassen, B. (2002). *The Empirical Stance*. New Haven: Yale University Press.

Wallace, D. (2006). In defence of naivete: The conceptual status of Lagrangian QFT. *Synthese*, 151: 33–80.

——(2012). *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. Oxford: Oxford University Press.

Weinberg, S. (1992). *Dreams of a Final Theory*. New York: Random House.

Wilson, M. (2006). *Wandering Significance*. Oxford: Oxford University Press.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. Translated by C. K. Ogden. London: Routledge.

7

What Can Contemporary Philosophy Learn from our 'Scientific Philosophy' Heritage?

Mark Wilson

> [I]t is so much easier to put an empty room tidy than a full one.
>
> W. K. Clifford (1879: 2)

I

Michael Friedman and I recommend that analytic philosophy should revisit the themes of its 'scientific philosophy' roots, maintaining that philosophical progress remains most secure when it has anchored itself firmly in real life dilemmas.[1] Through neglect of such ties, we judge that our discipline has slipped into a conceptual complacency from which it should be reawakened. One route to doing so is simply to think again about the methodological worries that troubled the pioneers of scientific philosophy, considered afresh in light of the gallons of discovery that have subsequently washed beneath knowledge's bridge. (Although few contemporary

[1] This exchange grew from a critique ('Logic, Mathematical Science and Twentieth-Century Philosophy'; henceforth LMS) that Michael Friedman composed for a symposium in *Noûs* on my *Wandering Significance* (Wilson 2006). I am deeply indebted to Friedman, Robert Brandom, Michael Liston, Anil Gupta, A. W. Carus, and Penelope Maddy for their very stimulating comments.

philosophers remember this, many of analytic philosophy's key doctrines were set long ago in response to various half-forgotten scientific crises arising in the late nineteenth century.) However, in surface rhetoric at least, the conclusions Friedman and I reach diverge substantially. There is probably less to these disagreements than meets the eye but they stem from different appraisals of where the true merits of our 'scientific philosophy' heritage lie. I fancy that a wider audience may find a brisk survey of the salient issues useful, in order that analytic philosophy might once again serve as a trenchant 'critic of concepts' in the mode of our great philosophical forebears. In the main, I believe that this revivalist project requires that we think again about the central place that *logic* currently occupies within our thinking.

To this end, I will compare Friedman's recent *The Dynamics of Reason* (DofR) with my own *Wandering Significance* (WS), supplemented by Friedman's subsequent essay 'Logic, Mathematical Science and Twentieth-Century Philosophy' (LMS, 2010a).[2] DofR notably criticizes present day philosophy for its conceptual timidity, observing that it rarely serves as the fierce *enemy of ideological complacency* that it did back in the glory days of Helmholtz, Mach, Poincaré and Einstein and their positivist followers, Carnap, Schlick, and Reichenbach.[3] In fact, a fundamental mission of Friedman's LMS is to remind us of the strong 'anti-classical' credentials of the entire scientific philosophy movement, where 'anti-classical' in this context designates any approach to 'conceptual content' that opposes traditionalist presumptions that such 'contents' can be readily surveyed by mental introspection—WS cites Bertrand Russell's *Problems of Philosophy* as a paragon of such 'classicism.' As Friedman correctly stresses, strong strains of resistance to 'classical conceptual thinking' arose far earlier than Russell's 1912 book, despite WS's apparent suggestion that this revolt commenced thereafter. But the latter impression traces merely to my expositional ineptitude—it was not intended. (I hadn't attempted a faithful historical commentary at all.) Indeed, it can be plausibly argued that reservations about 'concepts classically conceived' lay implicit in the early battles over the suitability of Newton's form of mechanics. WS cites Russell's exposition centrally only because I feel that he marked out the vital contours of

[2] The present essay is a slightly altered version of the reply I offered to Friedman in that volume. The second essay included here reflects conversations that I've subsequently had with Friedman and others, in Birmingham and elsewhere.

[3] For want of a better concise term, I shall loosely lump all of these thinkers together as 'positivists.'

the classical tradition in an especially admirable way (and, at the same time, indicated how many standard anti-classical gambits could be plausibly co-opted under its banner). But most of the vital strands within anti-classicism were fully on the field before then. And I have always viewed the positivists themselves as comrades in the struggle against 'classical concepts.'[4]

Nonetheless, these thinkers didn't make much of an appearance within my narrative because I feel that they had pursued the *wrong branch* within historical anti-classicism.[5] Specifically, they favored a view of 'conceptual content' that relies strongly upon *logical structure* in the manner that Friedman deftly outlines in LMS. Specifically, they believed that logic (including set theory) offers sufficient constructive tools to provide a general

[4] However, a complication on this score needs to be acknowledged, in deference to the historical record. It is one thing to *recognize*, as the scientific philosophy tradition plainly did, that excessive loyalty to 'classical conceptual content' poses barriers to the progress of science and quite another to explain exactly what is *wrong* with the notion in itself. The historical path of least resistance was simply to declare, 'Oh, such "content" exists alright, but science isn't obliged to *worry* about it.' This philosophical meekness leads my favorite avatars of 'scientific philosophy' (Helmholtz and Hertz) to declarations of the ilk: 'Science can only know the external world *qualitatively* up to isomorphism.' Plainly, such opinions endorse the coherence of a strongly 'classical' view of content through their tacit understanding of 'qualitatively.' For myself, I concur with vociferous naïve realists in finding such 'veil of perception' doctrines repugnant, but I am not willing to hypostatize some mythical 'common-sense world' simply so that we might readily stay in touch with it. No; the universe in which we actually dwell is a quite strange place and any acceptable 'realism' must warmly acknowledge that unhappy fact, I warrant.

In fact, the best critics of over-inflation and rigidification within our naïve understanding of 'content' are to be found amongst British thinkers such as Thomas Reid and J. L. Austin, although they rarely appreciated the 'barriers to progress' worries that correctly pushed the scientific philosophy tradition towards more radicalized conclusions overall. Indeed, the structure of WS is best viewed as a self-conscious effort to follow the lead of Hertz and Helmholtz while employing some of the diagnostic tools pioneered by Reid and Austin. But fulfilling that program is a tall order, for it requires that we mollify allied misapprehensions about 'perceptual content' in a plausible manner as well. Doing all of this requires a vast amount of subtle diagnostic work and I do not pretend to have carried out all of the designated tasks in a satisfactory manner. But one of the reasons that WS grew so bulky is that it strived to warn my fellow 'scientific philosophers': 'No, you can't simply *ignore* "classical content"; you must also *diffuse* the mythologies of exaggeration upon which it deeply depends.'

[5] Since I wrote the above, a helpful correspondence with André Carus has led me to look more closely at Carnap's pronouncements on 'meanings' and 'concepts.' In my assessment, they are surprisingly perfunctory and unhelpful, especially in light of the deeper critiques offered by earlier thinkers such as Mach, Helmholtz, and Schlick. This critical inadequacy only underscores my belief that Carnap should not be regarded as a wholly optimal representative of my 'scientific philosophy tradition.' (Heretofore, I have tended to read Carnap through the lens of Quine's doctrines, but I now believe that I improperly diminished the latter's originality in crediting such themes to Carnap.) An exchange between Carus and myself on these topics can be found in 'Carnap's Ideal of Explication and Naturalism' (Wagner, 2011).

framework for all (anti-classical) conceptual construction, through 'implicit definability' in the arms of a strong, logically interlaced 'theory.' This presumption, although reasonable at the time, left the positivists 'barking up the wrong gum tree,' in Austin's phrase, and the after-effects of this wrong turning strongly affect analytic philosophy to this day. In contrast, I discern a quieter and gentler anti-classical pattern within the historical record that promises a more satisfactory template for a revived 'scientific philosophy.' Within this alternative approach, logic proves decidedly subservient to other forms of inferential procedure.

Before we inspect these two patterns within scientific anti-classicism, let me reiterate a central theme within DofR: the varying recommendations of the scientific philosophy movement grew out of a fundamental quest for greater *conceptual liberty* within science. By the mid nineteenth century, a flurry of strong winds had begun reshaping mathematics and physics in radical and unexpected directions.[6] A paradigm of such pressures can be found in Friedman's central example: shifts in attitude and inferential technique within geometry. Plainly, some novel strain of philosophical thinking was required to make sense of it all, for naïve, traditionalist assumptions about 'conceptual content' were apt to resist such advances. The watchword of the time was that the 'free creativity of the scientist' needs to be warmly respected in our liberal toleration of conceptual ingredients. But what, exactly, might that phrase mean?

We moderns have entirely embraced this 'free creativity,' in that protracted experience has taught us to accept whatever strange conceptual obligations Nature lays upon our plates. Indeed, perhaps we have become *too complacent* in our forbearance, for we are apt to forget that, in real life circumstances, acting in a truly open-minded fashion is not as easy as one might fancy in the armchair. (It is much harder to renounce ingrained bigotry in practice than in theory.) Repetitious episodes in the history of science remind us of these conceptual roadblocks: crucial notions are commonly greeted with bafflement and resistance when first proposed. Why does this continue to occur? DofR observes that the positivists provided the beginnings of a plausible answer to this puzzle in their treatment of the 'relativized a priori' content and contends that we should not relinquish this explanatory beachhead lightly in our own thinking. Yet

[6] An allied push towards 'free creativity' arose within pure mathematics as well. For a brisk survey, see Wilson (forthcoming).

something gone awry within contemporary philosophy has persuaded it to do precisely that: we forget or minimize the conceptual difficulties of real life practice and allow the radical remedies they require to slip from view. In an original and distinctive stretch of diagnosis, Friedman traces much of this amnesia to the untoward influence of Quine's writings, a theme I will amplify in the next section.

Central to Friedman's argument are the great geometrical *thought experiments* suggested by Helmholtz and company, as these provided the *crucial semantical pivot* that prepared a path for Einstein's later Relativity revolution. Helmholtz, for example, told a 'Flatland' story of 'what one would see' living in a world of constant but mild curvature obedient to a natural adaptation of traditional geometrical optics (despite the fact that he knew light was actually a wave phenomenon of some kind). But how can *thought experiment whimsies* play such important roles within conceptual development? The positivists' response to this mystery, as reconstructed by Friedman, will be the subject of section II. His crucial observation is that, if contemporary philosophy is to resume its proper calling as a 'critic of concepts', it must develop some parallel account that renders justice to the peculiar semantic effectiveness of such 'thought experiment' pivots. In fact, if we fail to attend to this chore, then we probably lack a plausible answer as to why philosophy is useful at all. The conceptual transitions facilitated by the musings of Helmholtz et al. should be regarded as emblematic of our fundamental intellectual obligations, as they provide an especially convincing illustration of how overtly 'philosophical thinking' aids intellectual progress. But current complacency about the breadth of our intellectual grasp (e.g. in the guise of glib appeals to 'all logically possible worlds') fancies that such developmental problems lie entirely in the past and that philosophy has no further role to play in such matters.

I agree with Friedman on all of this, but wonder if the old positivist explanation of conceptual resistance can't be reworked in a different manner than he suggests. In this essay, I will employ my alternative diagnosis to illustrate some of the key theses about 'semantics' advanced within WS.

II

As noted above, the central impulse driving the rise of scientific philosophy grew from a desire to liberate the 'free creativity of the scientist' from what

Gauss dubbed the nattering 'complaints of the Boethians.' But what philosophical conception of 'concept' can underpin such an ideological liberalism? By the beginning of the twentieth century, it became widely assumed that scientific concepts could be supplied with adequate 'meanings' simply through an inclusion within a properly articulated theoretical frame. (In the standard jargon, a logically articulated theory will *implicitly define* its embedded terminology.) In its crudest formulations, this 'organizational' point of view claims that, in the final analysis, all that's really vital for a 'conceptual system' is that it lay down reasoning pathways able to convey us from correct input measurement numbers to correct output numbers. As Richard Feynmann puts the matter: 'The whole purpose of physics is to find a number, with decimal points, etc.! Otherwise you haven't done anything' (Manin, 1981: 35).[7] From this point of view, traditional worries about conceptual opacity or unintelligibility can be generically dismissed with an airy: 'Hell, we don't really need to "understand" nature in the manner you expect; we merely need to predict its courses accurately.' Hertz, in a celebrated passage from his *Principles of Mechanics*, articulates the recommended policy of conceptual tolerance more delicately as follows:

> The most direct, and in a sense the most important, problem which our conscious knowledge of nature should enable us to solve is the anticipation of future events, so that we arrange our present affairs in accordance with such anticipation . . . In endeavoring thus to draw inferences as to the future from the past, we always adopt the following process. We form for ourselves images or symbols of external objects; and the form which we give them is such that the necessary consequents of the images in thought are always the images of the necessary consequents in nature of the things pictured . . . [W]e can [thus] in a short time develop by means of them, as by means of models, the consequences which in the external world only arise in a comparatively long time. (Hertz 1952: 1)

He correctly observes that many strands commonly utilized within mechanics do not harmonize with one another perfectly from a strict logical point of view:

[7] Manin himself comments: 'This is an overstatement. The principal aim of physical theories is understanding' (1981: 35). Some of Manin's reasons for this counterclaim will emerge in part III of this essay. To avoid potential misunderstanding, let me stress that few 'scientific philosophers' have fully embraced the crudely operationalist themes that Feynman sounds in this passage. (Indeed, as indicated in note 3, it is often unclear what their developed 'semantic' views might be.) I employ this quotation as merely the coarse expression of an inclination that generally assumes subtler forms.

> [W]e have accumulated around the terms 'force' and 'electricity' more relations than can be completely reconciled amongst themselves. We have an obscure feeling of this and want to have things cleared up. Our confused wish finds expression in the confused question as to the nature of force and electricity. But the answer we want is not really an answer to this question. It is not by finding out more and fresh relations and connections that it can be answered; but by removing the contradictions existing between those already known, and perhaps by reducing their number. When these painful contradictions are removed, the question as to the nature of force will not have been answered; but our minds, no longer vexed, will cease to ask illegitimate questions. (Hertz, 1952: 8)

Note that Hertz's sole complaint is that *too many* conceptual stipulations have been laid down in mechanics to cohere with one another logically.[8] In his final sentence, he stresses the fact that he is *not* criticizing the notion of 'force' on the traditional grounds that its 'content' seems 'occult' or 'non-understandable'; he plainly rejects such classical criticisms as ill suited to the 'free creativity' of science. He sees his task as entirely one of reducing mechanics' excessive collection of inferential principles to logical consistency without jettisoning important derived principles such as the conservation of energy along the way. It is worth observing that Wittgenstein chose Hertz's final sentence as the projected motto for his *Philosophical Investigations*, a point to which I'll return.[9]

Few steps are required to pass from Hertz's point of view to the *logic-centered picture* of concepts favored by many in the positivist movement of the twentieth century and their sympathizers, where 'conceptual content' becomes explicitly elucidated as *implicit definability* within a self-consistent encompassing structure. The anti-classical approach to 'conceptual content' that this doctrine facilitates explains why this school came to regard a well-articulated logical structure in the sense of *Principia Mathematica* as the sine qua non of the theoretical frame required.[10] (Indeed, a maximized

[8] For example, Max Jammer (1962) completely mistakes Hertz' motivations. For a better account, see Jesper Lutzen (2005). David Hilbert makes very similar remarks when he explains why he has set 'axiomatizing physics' on his famous set of problems that mathematicians should address during the twentieth century. See Felix Browder, (eds.) (1983). Hilbert's thoughts on axiomization and 'implicit definability' influenced the positivists greatly.

[9] Cf. Baker and Hacker (2005: X). I first learned this salient tidbit from Michael Kremer.

[10] Of course, twentieth-century conceptual classicists generally relied upon the same logical formalism as well, although they didn't employ the framework as a source of 'distributed normativity' in the sense of WS. In his more 'structuralist' moments, Russell often comes close to an 'implicit definability' point of view with respect to the predicates of physics.

liberalism suggests that these logical benchmarks should represent the *only substantive restrictions* upon scientific conceptualization.) Under the banner of these liberal and deeply anti-classical attitudes, logic came to assume the 'distinctive role' that Friedman ably emphasizes in his LMS discussion of our twentieth-century forebears. Once the required structural ties have been logically forged, the central task remaining to conceptualization is to indicate how such linguistic provisos can be fruitfully 'coordinated' with empirical reality. It then becomes natural to expect that a modest range of structural provisos must be activated before *it even makes sense* to speak of the 'empirical predictions' available within a specific framework. Sundry 'prime the pump' preliminaries must be set in place before the formalism can draw empirical water. (Plainly, prerequisites of this type reflect a deeply Kantian understanding of how 'concepts' function.) Such framework-supporting presumptions constituted the 'relativized a priori' that Carnap and the early Reichenbach substituted for Kant's more absolutist architectural requirements.

At first blush, it might seem as if one could plausibly retain an 'implicit definability' story of conceptual content while simultaneously rejecting any clean relativized a priori, on the grounds that developmental pressures are apt to quickly muddy what qualifies as 'presupposition' and what qualifies as 'empirical after-effect' within science. Indeed, this 'first blush' alternative exactly represents Quine's perspective on such matters. But Friedman maintains that this blithe abandonment of the 'relativized a priori' comes at a severe intellectual cost, for it cavalierly abandons a sophisticated understanding of philosophy's proper role as a useful 'critic of concepts.' He astutely observes:

> [Modern analytic philosophy] cling[s] to the idea of a specially characteristically philosophical enterprise (of 'mere conceptual possibilities') inherited from the . . . analytic tradition, while simultaneously ignoring completely the distinctive conception of logic which made this tradition possible in the first place. (LMS, p. 542)[11]

Here Friedman captures a theme that WS likewise explores: much contemporary thinking on 'concepts' represents a hodge-podge of themes extracted from a philosophical past that was more disciplined within its

[11] To avoid confusion, I've omitted the adjective 'classical' before the phrase 'analytic tradition' only because Friedman employs the term differently from me, in that he means to embrace the entire positivist movement within his sweep, whereas I would classify most of it as pursing *anti*-classical conceptual goals.

own thinking. Many readers of WS have come away with the impression that it accuses contemporary figures such as David Lewis, Saul Kripke, Jerry Fodor et al. of being 'classicists,' when they plainly resist such a classification along several fronts. In truth, to consider them 'classicists' represents a gentler appraisal of their proposals than I actually entertain, for I largely discern an ill-sorted mélange of classical and anti-classical themes that have descended, without proper scrutiny, from earlier philosophical traditions. And Friedman clearly believes the same thing.

However, he is concerned with a deeper point as well:

> It is especially ironic that what Wilson calls Quine's 'hazy holism' now gives aid and comfort to the conceptual ambitions of contemporary analytic philosophy by allowing us to imagine that our armchair philosophizing merely occupies an especially central and abstract level in our total empirical theory of the world—and, as such, it operates within the very same constraints, of overall 'simplicity' and 'explanatory power,' governing ordinary empirical theorizing. In this way, the great opponent of the analytic/synthetic distinction unwittingly made room for essentially a priori philosophizing through the back door. (LMS, p. 544, footnote 17)[12]

The claim is that, in casting aside the positivists' notion of 'semantic prerequisite' with the unwanted bath waters of the analytic/synthetic distinction, modern philosophy has relinquished the vital tools it requires to understand the *peculiar difficulties of real life conceptual advance*. Without some tangible foundation within a 'relativized a priori' (or something like it), all 'concepts' begin to look essentially alike, allowing their contemporary student to fancy that she no longer needs to probe the extensive annals of surprising adjustment within science before framing apt conclusions about 'how concepts work.' But in such generality folly lies, for one can truly appreciate why we *need to worry about concepts* only through closely examining issues such as the underlying causes of ideological resistance within science. DofR hopes that the great thought experiments of nineteenth-century geometry will inspire us to become genuine conceptual critics again, yet our contemporary scholar sees nothing significantly 'semantic' in any of it—just a bunch of guys shifting the meanings of 'distance' and 'time lapse.' She will dismiss all of Friedman's historical nitty-gritty with an insouciant, 'So stuff happens; let's get back to concentrating upon DOG and

[12] Of course, the congruent legacy of the British 'common-sense' school of the 1930's also encouraged a blurring of traditions through an improper minimization of the conceptual challenges presented within real life scientific practice.

DOORKNOB.' And Quine's influence, Friedman claims, has laid the groundwork for such detachment from real life struggle.

As previously suggested, Friedman's rejoinder emphasizes the need to explain why preliminary thought experiments so often prove essential to conceptual toleration: how can *patent fictions* prove essential prerequisites to proper intellectual grasp? He reminds us that the positivists' approach to conceptual innovation emerged precisely from their efforts to render the semantic benefits conferred within the great nineteenth-century discussions of non-Euclidean geometry into a formal and adequately diagnosed frame. This codification relies upon both the 'implicit definability' approach to conceptual content (to allow maximal 'free creativity' in science) and the isolation of a 'relativized a priori' as presuppositional core (to explain the basic manner in which the new constructs 'coordinate' with empirical observations). With respect to the latter, merely grasping the purely mathematical structure of a novel set of notions does not show how they can be successfully arranged around the familiar verities of measurement technique. (Qua mathematical structure, I understand the complexified projective plane well enough, but I don't see how I could *live* in such a joint.) Clearly, the great geometrical thought experiments provided this crucial bridging service: they adequately illustrate how such 'coordinative' resettlement of familiar experience within a non-Euclidean frame can become accomplished. As instruments of conceptual advance, these patently 'philosophical' constructions provided a necessary prerequisite for the eventual development of general relativity, despite the fact that the latter eventually adopted more complex strategies for employing non-Euclidean concepts than were exemplified within the original thought experiments. Without a mollifying thought experiment bridge, conservative parties had no easy pathway to appreciating the novel 'possibilities' offered within general relativity's unfamiliar frame. For the positivists, their 'relativized a priori' merely represented the formalized treatment of the fundamental mathematics-to-experiment ties required to pull non-Euclidean skeptics into richer conceptual pastures.

Accordingly, the positivists' story provides an attractive portrait of a specific manner in which philosophy can serve as a useful 'critic of concepts,' opening up unexpected conceptual terrains in advance of knowing whether Nature will fondly embrace such semantic recalibrations or not.

To press these positivist advantages upon us, Friedman must first rectify a prevailing interpretative injustice. Contrary to Thomas Kuhn's asseverations

in the *Structure of Scientific Revolutions*, the positivists were fully aware that science is not invariably 'cumulative' in character and, as we've seen, much of their thinking was strongly shaped by a desire to *diagnose* the underlying character of typical bumpy resistance to 'scientific revolutions.'[13] Indeed, they appreciated the crucial role that thought experiments play in making the classical-physics-to-relativity transition appear *rationally adjudicated*, rather than merely representing the crude triumph of one doctrinal prejudice over another in the 'political' manner that sociologists of science often suggest. As Friedman perceptively stresses, the 'relativized a priori' captures the *conceptual prerequisites of coordination* that a critic must grasp before she can think of herself as 'possibly living within a non-Euclidean world.'

Pace Quine and his followers, modern sanguinity with respect to 'conceptual content' provides no comparable means for explaining such common semantic patterns within scientific development, despite their ubiquity:

> Our best current historiography of science requires us to draw a fundamental distinction between constitutive principles, on the one side, and properly empirical laws formulated against the background of such principles, on the other. (DofR, 43)[14]

In so doing, recent philosophy has abandoned the diagnostic tools required to operate as effective critics of conceptual complacency in the manner characteristic of the entire 'scientific philosophy' movement. Worse yet, we have deluded ourselves that this noble calling 'no longer represents

[13] In truth, 'explaining the utility of the thought experiments' was just one reed in the positivists' muddle of motivations; it is the singular merit of DofR to have brought forth this submerged theme clearly. All the same, why did Kuhn so thoroughly mistake the positivists re 'accumulationism'? I am not certain, but the misreading may derive from the writings of mid-century 'logical empiricists' such as Ernest Nagel. The latter's best work lay in the developmental history of mathematics (*Teleology Revisited*) and his own essays emphasized some of the same 'Neurath's boat' themes to be highlighted in section III of this essay.

[14] With the phrase 'best history,' I believe Friedman is being too generous, for we should entertain strong reservations about 'histories' that regularly omit the nitty-gritty details of science (such as the conflicting mathematical entanglements that bothered Hertz) in favor of the attendant political and ideological squabbles. Studies in the sociological manner of Kuhn and his followers have frequently amplified the latter to the point of absurdity, while ignoring the commonplace inferential mysteries that should rightly leave any rational agent perplexed and uncertain where to turn. For philosophy's purposes, our 'current historiography of science' strikes me as far too lopsided to be embraced wholeheartedly, unless the term also encompasses the excellent 'amateur histories' regularly penned by retired scientists in their leisure years.

philosophy's affair.' It is as if we were only willing to study meta-ethics and had abandoned all normative concern to the preachers and politicians.

Once again I agree with much of this and share with Friedman an unbounded enthusiasm for the peculiar *semantic virtues* offered within the great geometrical thought experiments. Nonetheless, I don't believe that the positivists did a very good job in explicating why such proposals should qualify as 'semantical' and in the next section I shall sketch an alternative diagnosis of their special character. Much of the problem traces to the well-known fact that, as soon as one attempts to identify a presuppositional core within any concrete set of scientific doctrines, one quickly becomes stymied, for few of the 'philosophical' remarks attached to the sundry thought experiments that inspired Einstein's advances actually became uncontroversially implemented within his finalized proposals. Nor are the 'philosophical virtues' commonly consigned to Einstein's advances cleanly displayed within his completed theories either. For example, positivist-influenced commentators regularly claim that Special Relativity deftly avoids the extraneous aether/matter ontology that troubled Lorentz. Perhaps, but Einstein's formal shift encouraged a point of view where matter should be modeled by point particles situated within an electrically active vacuum. As soon as any charged particle of this type accelerates, one runs into the same horrible mathematical divergences that Lorentz attempted to cure by other means. (*Any* manner of matter/field pairing is apt to prove just as mathematically troubling as the aether/matter coupling.) Likewise, general relativity is often credited with 'explaining' gravitational activity as freely falling matter following geodesics. Yet Einstein's finalized equations demand that its 'matter' become continuously distributed in a manner that makes 'freely falling matter follows geodesics' mean scarcely more than 'geodesics lie on geodesics.' The resulting mélange of mush and avoidance renders the clean 'semantic prerequisites' of the positivists vulnerable to Quinean objections.

The generic root of such difficulties traces to the brute fact that science rarely advances as a complete and coherent 'framework' in any reasonable sense of the term. (It 'comes at us in sections,' to quote Fred Astaire in *The Band Wagon*.) At present, there is no happy method for wedding Maxwell's electromagnetism or general relativity to a fully embodied 'theory of matter' within wholly classical circumstances (let alone to quantum theory). Practitioners typically tiptoe around these alignment problems by modeling matter's influence opportunistically as 'boundary conditions' or

'transmission coefficients' in manners that plainly cannot capture the general situation. In brute fact, what Einstein actually contributed to science was a motley collection of equations and inferential techniques that clearly 'have some truth in them' but don't cohere with one another perfectly and also leave huge swatches of physical behavior quite unspecified. (None of this should be regarded as belittling his great achievements at all; it is simply a comment upon the typical character of physical advance.) Of course, the fact that classical mechanics had evolved its 'doctrines' in this haphazard and opportunistic manner was the source of the conceptual conflicts that bothered Hertz.

The positivists were forced to freeze assumptions about ordinary matter (e.g. 'measurement rods' on a macroscopic scale) into artificial fixity because these prerequisites needed to serve as the *stable core* around which the 'free creativity of the scientist' can liberally range.[15] These doctrinal requirements have the unfortunate side effect of locking conceptual advance in celestial mechanics to presumptions about the behaviors of everyday matter on a localized scale. Yet no science yet devised has managed to approach both tasks in a truly harmonious manner.[16] Such are the familiar problems attaching to the 'relativized a priori' and 'coordinative definitions' of the positivists. Friedman is, of course, fully aware of these woes, but he advises us to not lightly abandon our tasks as conceptual critics thereby. Furthermore, in recent work, he has offered an increasingly sophisticated set of proposals for repairing these deficits from an essentially Kantian perspective. Still I am not convinced, partially because I don't understand the proposals well enough to sort out, at any historical moment, what should qualify as a pre-empirical 'presupposition' within a science. (Friedman cites the geometry of inertial frames as an example within 'classical mechanics' but the messy subject called 'classical mechanics' in

[15] Actually, the early positivists mostly tried to employ phenomenal verities as the core to be organized, but this position enshrines improper psychological description just as warmly as the 'measurement rod' approach locks in physical misdescription. In his 'Replies' in the festshrift *Discourse on a New Method*, Friedman appeals to the 'light principle' as a means of evading some of these problems, although allied problems attach to 'light' insofar as we interpret the latter phrase as signifying an identifiable form of physical emission.

[16] Such woes remind me of the unhappy fates of the hippie communes of the 1960s that, in striving for maximal liberty along several dimensions of conduct, failed to incorporate the subtle controls and tolerations that rescue more mature societies from the vicissitudes of human character. Indeed, this analogy essentially epitomizes the entire argument of section III in a nutshell!

actual practice contains many strains that do not rest unambiguously upon this foundation.) Of course, the fact that real life science often contains 'untidy rooms' stuffed with furniture that stylistically clash with one another represents the deepest source of the Quinean conviction that we can repair its problems only in a gradual, 'Neurath's boat' manner that is unlikely to reveal any standing 'presuppositions.'[17] However, lacking an alternate understanding of the bridging capacities offered within the prototypical thought experiments, it becomes hard to summon philosophy to its critical duties on Freidman's model, as long as it appears likely that the only 'semantic prerequisites' philosophers will plausibly codify are destined to wind up *rejected or ignored* within the very theories for which they allegedly serve as 'prerequisites.'

In LMS, Friedman deftly recounts how twentieth-century analytic philosophy, from both 'classical' and 'anti-classical' perspectives, commonly assigned logic and set theory a very central role within their various accounts of 'concept formation.' It is a characteristic theme within WS that this logistic emphasis was, in fact, a mistake and that we should instead attend more closely to other inferential invariants in our thinking about the 'core contents' of critical concepts. (Set theory has an important role to play, but it is not that assigned within the story that Friedman recapitulates.) Although I had not thought of these issues when I wrote my book, it strikes me that Friedman's vital insistence upon the *semantic utility* of the great geometrical thought experiments might be accommodated within WS's framework without any requirement that we clean up science's messy stables in the thoroughly hygienic manner of the positivists.

III

As indicated in an earlier footnote, I generally find histories of sciences pursued in Kuhn's manner to be rather one-sided, for they typically ignore the gentler forms of historical adjustment within physics and mathematics that often prove crucial to the 'semantic' issues at issue. Here I have in

[17] Although WS frames conclusions that sound more 'Quinean' than the neo-Kantian position espoused within DofR, it reaches those conclusions along decidedly non-Quinean paths. Or so it strikes me. Although I would never wish to diminish his salutary influence upon my thinking, I am troubled that much of Quine's actual argumentation rests upon the same *mythologies of scientific methodology* that severely handicap philosophical progress to this day. In this appraisal, I am fully in agreement with Friedman.

mind the gradual shifting of textbook 'definitions' and the highlighting of fresh mathematical tools ('invariant', 'potential') that occur upon an almost daily basis within a healthy discipline, generally unfolding in a quiet and gradualist manner. Despite the near invisibility of these shifts, they frequently exemplify probing 'critiques of concepts' as philosophically meritorious as those exemplified within splashier episodes such as the non-Euclidean revolution. To appreciate the importance of these adjustments, let us look more closely at the inferential difficulties that commonly trouble working science.

In particular, let us attend to the reliability of *localized reasoning methods*, such as the mathematical techniques that a physicist or engineer employs to extract nuggets of useful conclusion from otherwise recalcitrant equations (to 'get the numbers' that Feynman wants). And here we must observe that the most powerful inferential schemes utilized within applied mathematics often prove erratic in their performance: they sometimes work well and sometimes work badly, without displaying evident marks to distinguish the cases. Allied woes are rarely evident within the well-behaved inferential patterns of first order logic, whose atypically cooperative behavior has bewitched many philosophers into overlooking the computational logjams that substantially shape the topography of working science. Indeed, my main complaint about twentieth-century positivism is that it has inadvertently encouraged this *logic-entranced obliviousness* to the substantial hazards of real life inference, a condition from which contemporary philosophy has not yet fully recovered.

As a prototypical case of erratic inferential behavior, Cauchy's worries about 'rigor' began in a puzzlement why widely employed series expansion techniques in celestial mechanics allow us to solve Kepler's formula $\mathbf{M = E - e \sin E}$ correctly over certain ranges for $\mathbf{e}$ but not over others. What hallmarks must a reasoning principle further display if it is to be justly trusted? Understanding such issues often requires that the mathematician provide what WS calls a 'picture of our computational place in nature'. That is, they construct a model that welds together the reasoning pattern R and worldly behaviors C to which the inferential pattern allegedly answers. Our mathematicians then investigate how ably R manages to relate the membership of C within this generic setting. We rarely know much concrete detail about C on an a priori basis: we are, after all, hoping to establish that the conclusions we can reach by applying R to the starting data **D** will prove *trustworthy in any conceivable circumstances* in which we

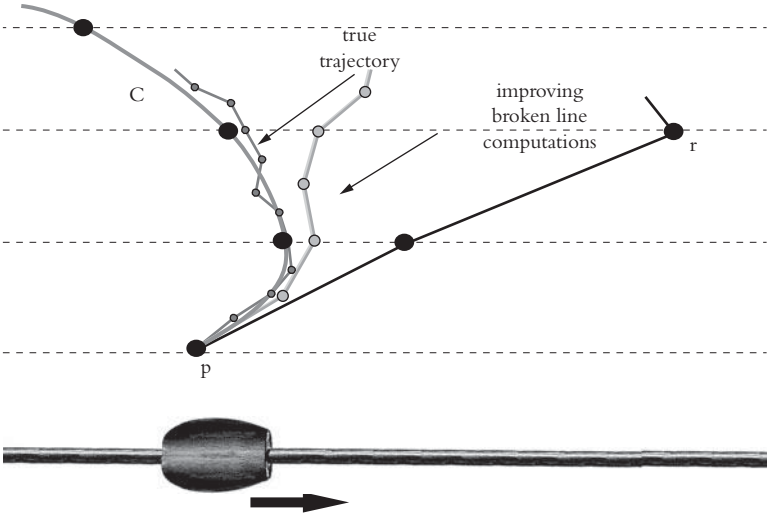

Figure 7.1 Computations converge to correct trajectory.

imagine R might be applied. So mathematicians frame the relevant possibilities in huge classes C wide enough to model any feasible manner in which the real world might make the data **D** true.

Let me illustrate such a 'computational portrait' in a bit more detail.[18] Most basic physical processes are governed by differential equations (or something very much like them). As a simple case in point, suppose that an ordinary differential equation (call it 'E') of the form $dx/dt = f(x)$ governs how a bead slides along a wire. As such, the bead must travel along some unknown curve **C** within the general class of continuous curves C. Our formula E states that the velocity of the bead is linked at every temporal moment to its current position by a simple f-rule (whenever the bead reaches the point $\underline{p}$, its current velocity must be $f(\underline{p})$). As such, E only fixes how the bead behaves *at each infinitesimal moment,* it doesn't directly tell us how the bead will move over any longer time interval. We now hope to find some reasoning rule R that starts with the equation E and $\underline{p}$'s location as its starting data **D** and reasons to reliable conclusions about how the

[18] A remark developed more fully in ch. 4, §iv of WS is this: if we convert Hertz's metaphor about 'images' into concrete computational terms, the closest fit we are likely to find is a 'marching scheme' of Euler's method type.

bead will behave at some later time $t^{\star}$. But don't be misled by the exact solutions one is taught in freshman calculus; such manipulative techniques only rarely work. If we want general purpose reasoning methods, we need to look at so-called 'numerical schemes' of the ilk commonly installed within computer programs. A basic prototype of such a policy is *Euler's method*. Indeed, its workings are quite intuitive. If the $\underline{f}$ in our equation is readily computable, we can divide the bead's temporal development into short time steps t and argue that its velocity at the starting time t_0 should supply a *plausible estimate* for where the bead will later land at time $t_0 + t$. That is, starting at $\underline{p}$ and t_0, Euler's method instructs us to draw a straight line leading to the point $\underline{q}$ where the bead would land if it had maintained a *constant* $f(\underline{p})$ feet/sec velocity over the entire t interval. Applying Euler's procedure once again at $\underline{q}$, we next plot the point to which our bead would relocate at time $t_0 + 2t$ if it is assumed to maintain a constant $f(\underline{q})$ velocity after it passes point $\underline{q}$. Repeating this process, we obtain a broken line graph $\mathbf{C}^{\mathbf{E}}$ which, if the reasoning method is sound, should lie fairly close to the unknown true solution **C**, at least for a reasonable span of time.

But is Euler's rule genuinely reliable? Without further restrictions on E, we can immediately see that its reasonings must prove *inherently fallible*, for suppose that E supplies the bead with a sharp kick *inside* the first t interval (it is easy to invent a f term that will do this). Rather than remaining near the bead's true **C** location $\underline{q}$, the $\mathbf{C}^{\mathbf{E}}$ we calculate will mistakenly augur that the bead travels to a completely erroneous position $\underline{q}^{\star}$. Such discrepancies between predicted and actual landing places generally become more severe as the plotting steps are repeated, possibly eventuating in a broken line plot $\mathbf{C}^{\mathbf{E}}$ quite unlike the bead's actual trajectory **C**. This is no idle worry, for if one is not careful in real life computing, the blind application of even the best computer routines will cheerfully plot alleged 'solutions' that bear no resemblance to the true curves carved out in Mathematicsland by the target differential equations.[19] Obviously, such difficulties trace to the fact that we finite humans can only inspect how **C** behaves at staggered

[19] For example, modern mathematicians have framed an impressive set of conclusions with respect to the 'chaotic behavior' of various differential equations, but these conclusions have been generally reached by computer techniques of the fallible sort just sketched. There remains an outside chance that all of these results represent spurious artifacts of the numerical methods utilized.

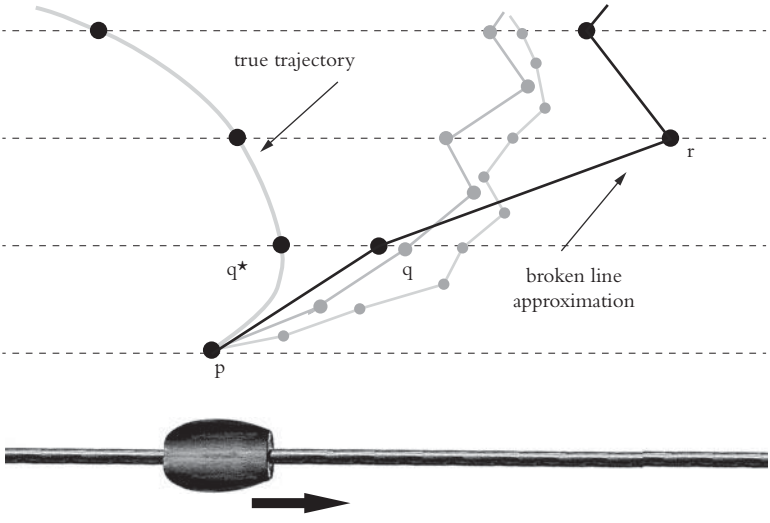

Figure 7.2

t intervals, allowing sufficient wiggle room for equation E, working at an infinitesimal scale, to decide to do something utterly different to our bead in the moments when we're not 'looking' (i.e. calculating). We must therefore hope that we can find some way to restrict the E's to which Euler's method is applied so that our calculated $\mathbf{C}^{\mathbf{E}}$'s will stay true to the curves **C** in C that meet additional provisos **M**. As it happens, there are a range of restrictions on the f's permitted in E that will turn this trick. But establishing firm 'correctness' results of this type is often rather hard and commonly the results fall short of the assurances that practicing physicists would like to see. (A substantial amount of scientific reasoning advances in the hazy twilight where the relevant R's are neither known to be sound or unsound.)

In weighing the reliability of Euler's rule within this C-based modeling, I hope that the philosophical reader recognizes that we are considering a non-logical equivalent to what logicians call a *soundness question*[20] with respect to a logical reasoning rule R: beginning with premises Γ ($\approx$ our data **D**), will the conclusions reached by **R** remain true in every model $\mathbf{M} \in \mathbf{M}$ ($\approx \mathbf{C} \in \mathbf{C}$) satisfying Γ? R is said to be *semantically justified* if it proves

[20] Students of numerical reasoning usually dub these as 'correctness results.'

'sound' or 'correct' under such a generic correlational study. A chief purpose of a 'semantic portrait' of descriptive language is to provide the underpinnings for 'soundness' evaluations such as this.

Shortly we'll investigate the manner in which important conceptual shifts are often associated with studies of this kind. However, let us first underscore a fundamental reason why we should expect that most rules of reliable reasoning will require non-trivial correlational scrutiny of this type. Inspecting the workings of Euler's rule, it is reasonable to feel: 'Gee, this rule clearly is operating *with the right idea*.' But why do we think that? Well, we'd like to presume that, if we could draw our straight line estimates over increasingly brief time steps t, the *complete collection* of broken line estimates $\mathbf{C}^{\mathbf{E}}$ we would frame will eventually surround the particle's true trajectory **C** as a cocooning envelope. Of course, only a god could truly trace the infinite mesh of estimates required in our hypothetical collection, forcing us, as mortal geometers, to terminate our labors at some finite step size t. As we saw, that enforced work stoppage potentially allows our E to supply an unexpected kick that our truncated reasoning process fails to notice. 'Close, but no cigar,' we might say, because such premature terminations make real life computers churn out absolutely dreadful inferential conclusions.

This gulf between our bead's real life curve **C** and the broken line estimate $\mathbf{C}^{\mathbf{E}}$ we can concretely calculate provides an emblematic illustration of our general *computational position within nature*: our actual inferential capacities commonly relate to real world behaviors in a *mildly transcendental fashion*, in the sense that fully deducing the bead's true physical circumstances demands the satisfaction of more conditions than we can inferentially check. In this regard, recall Richard Feynman's claim that 'The whole purpose of physics is to find a number, with decimal points, etc.' But our correlational studies warn us that, in the immortal words of Mick Jagger, 'you can't always get what you want.' Feynman may *want* those numbers, but there may be no computational path that will take him fully there. It is a brute fact that many natural processes run their courses in manners that lie aslant to any computational routine implementable by man or upon any of his swiftest machines, just as a ruler and compass alone cannot successfully trisect an angle. From a predictive point of view, this discrepancy between reasoning capacity and physical behavior is hardly ideal, yet it is surely preferable to living in a world whose unfolding processes relate to our calculations only in some horribly transcendental manner where we never get even *close* to a correct answer.

In such 'transcendental' circumstances, the need to consider inferential practice from a 'semantical point of view' becomes practically inevitable. Insofar as logic is concerned, this point of view is defended in influential works such as Michael Dummett's well-known 'The Justification of Deduction.' However, Dummett is far too *absolutist* in his point of view—a fault that arises, I think, from concentrating upon logical rules to the exclusion of the wide variety of other common inferential principles such as Euler's method. (These, after all, play a far larger role in the successful prosecution of everyday and scientific enterprises.) I explain what I mean by 'absolutist' in due course.

Returning to our central concerns, such C-based investigations often disclose that inferential schemes which qualify as unsound when judged by the lights of some received picture will prove correct if the correlational model studied is shifted to another interpretative basis ($C^\star$, say). We have noted that physicists commonly employ inferential principles in circumstances that the applied mathematician cannot semantically underwrite. If such successes occur repeatedly, it supplies a strong symptom that the *semantic underpinnings* of the reasoning method have not been rightly diagnosed. 'A method which leads to true results must have its logic,' advised the nineteenth-century mathematician Richard Woodhouse and this hidden 'logic' is often uncovered by setting old equations and inferential rules upon altered supportive moorings $C^\star$. (These shifts represent the quiet dimensions of conceptual reevaluation that gaudy Kuhnians generally ignore.) And it is not uncommon for a new reinterpretation to appear in hindsight as if it would have provided a *better reading* for the affected equations and data all along, if only earlier practitioners had possessed enough mathematical savvy to formulate the supplanting $C^\star$. For example, a large number of venerable physical equations dating to the nineteenth century have become reinterpreted within modern practice in quite sophisticated ways, usually through shifting the reading of the derivative operation 'd/dt' from its standard 'δ/ϵ' gloss to a smoothed over 'weak' reading or as a stochastic operator. Both of these replacements require considerably more technical machinery than the standard 'δ/ϵ' setup, but careful reflection upon physical principle commonly reveals that the adjusted readings better suit the physical circumstances to which the target equations have been long applied. For such reasons, practitioners often resist the presumption that they have 'changed the meanings' of the old equations: 'No, we didn't *chang*e their meaning; we finally got it *right*.'

In WS, I employ the evocative history of Oliver Heaviside's operational calculus to illustrate this tranquil vein of conceptual recalibration, but allied examples can be found everywhere in modern applied mathematics. Hans Lewy (who pioneered several important 'reinterpretations' of this ilk) remarked evocatively:

> [Mathematical analysis only supplies] hesitant statements . . . In some ways analysis is more like life than certain other parts of mathematics . . . There are some fields . . . where the statements are very clear, the hypotheses are clear, the conclusions are clear and the whole drift of the subject is also quite clear. But not in analysis. To present analysis in this form does violence to the subject, in my opinion. The conclusions are or should be considered temporary, also the conclusions, sort of temporary. As soon as you try to lay down exact conditions, you artificially restrict the subject. (As cited in Reid, 1991: 264)

'Analysis is like life'—I like that and would add—'and conceptual developments in every form of life commonly resembles those encountered in analysis.' Specifically, old words can be supplied novel supportive readings without undue fuss, for we frequently replace an implicit C modeling with a substantially altered C* without feeling that the terms' 'core meanings' have been substantially betrayed.

As such, these quiet episodes of semantic recalibration usually display a 'Neurath's boat' character, in that large blocks of successful inferential practice are maintained in place while the accompanying portrait C of their correlational underpinnings is adjusted, sometimes rather radically and in a manner that often opens fresh dominions to inferential exploitation in unexpected ways. And despite the apparently conservative character of the bootstrapping, the shackles of traditional constraints upon conceptual liberty can be cast off quite as effectively, if less dramatically, than through the blunt advocacy of a novel axiomatic frame. Indeed, 'definitional readjustments' in science generally display the same brutal indifference to established conceptions of 'conceptual content' as do the axiom-supported novelties favored by the positivists; the value of a proffered semantic picture C is adjudicated solely in terms of how ably it manages to capture the de facto correlational circumstances that allowed established usage to work as well as it has, no matter how poorly the prior practitioners of those practices appreciated these word-to-world supports. In such revisionary policies, we witness the same strongly 'anti-classical' attitudes to 'conceptual context' that the scientific philosophy movement has always championed, but less entangled with the constrictive holism

that makes the 'framework principles' of the positivists so hard to match with real life conceptual endeavor.

Over the course of the twentieth century, the conceptual conflicts within Newtonian mechanics that troubled Hertz became gradually resolved, not through finding an all-embracing axiomatic frame of the type he anticipated, but rather through *atomizing classical physics technique* into an accumulation of localized inferential gambits enjoying different semantic readings according to applicational context. The classical physicists in the nineteenth century had followed a variety of inferential signposts in blazing pathways to the celebrated modeling equations that modern engineers employ fruitfully in macroscopic circumstances even today. Figures like Hertz had presumed that these inferential conquests must rest upon some shared, if elusive, semantic foundation susceptible of axiomatic codification, but in making this assumption they had been (understandably) fooled by a common species of *semantic mimicry*, for rules R and R* can look very much alike, yet their word-to-world supports prove altogether different. Modern studies have shown that virtually all of the standard Victorian reasoning gambits were sound within the localized circumstances C where they were applied, but these supports C do not remain the same when we shift from one inferential success to another. So a proper resolution of the conflicts that troubled Hertz assumes a contextualized form: 'within applicational circumstances X, the proper setting for adjudicating the correctness of rule R should be C, but in applicational circumstances Y, it should instead be C*'.

In fact (although it is impossible to provide supportive argumentation here[21]), general theoretical considerations of an applied mathematics character suggest that frequent drifts in underlying semantic support should be commonly anticipated within well adjusted collections of *macroscopic variables*—viz. smallish clusters of descriptive vocabulary suitable for the very complex systems we encounter in everyday life. Our unavoidable observational and computational limitations force us to traffic in such 'reduced quantities' and their characteristic drifts in correlational support warn us to not hold such terms prissily to complete logical rectitude. The main intent of WS is to develop this line of thought in (probably excessive) detail.

[21] Wilson (2011) supplies a brisk recapitulation of the 'theoretical' side of this argument, but WS also endeavors to make its case through a lengthy array of case studies.

Accordingly, my own disinclination to cite the positivists as optimal exemplars of 'scientific philosophy' stems from my section (ii) complaints that their excessive 'implicit definability' demands require notions such as 'force' and 'rigid body' to lock together across broad domains when a proper resolution of the historical conflicts requires localized forms of strand-by-strand disentanglement. A symptom of the critical ills induced by such 'holism' can be found in the large and almost wholly misleading literature upon 'the problems of classical mechanics' spun out in the 1950s under the influence of positivist thinking.[22] Virtually every recommendation in such commentaries would have precluded advance in modern continuum mechanics had the practitioners of the subject paid any attention to them. Unfortunately, the same positivist-generated mythologies continue to adversely affect philosophical thinking today, even though the inspirational lights of the original program have largely dimmed.

In any event, it strikes me that the 'semantically adjusting role' of the classic geometrical thought experiments might be better accommodated within our gradualist story, which no longer demands that trustworthy 'framework principles' must be articulated within some coherent and all embracing theoretical frame. Although a proper analysis would take us too far afield, any investigation of the subtle ways in which inferential principles are actually weighed within real life applied mathematics reveals a number of ways in which *patently incomplete or simplified models* are commonly employed to 'try out' inferential procedures R in a bootstrapping fashion when one can't otherwise gain a reliable sense of how such reasoning techniques function over their intended domains. Thus applied mathematicians often develop a preliminary sense of how ably reasoning rule R operates over various *well-understood but inadequate models* $M^{\star}$ before they apply R to the subject matter in which they are actually interested. (Such behaviors will be presumably captured within some family of models M, but we can often learn very little about these M-structures until we begin to *trust* inferential tools like R.) But directly establishing the soundness of R over the desired M is often an intractable task, so our applied mathematicians satisfy themselves with checking R's soundness

[22] As noted above, Max Jammer's long series of 'The Concept of . . . .' primers epitomize these trends.

over the simpler M*.[23] Accordingly, gauging the 'contents' of key concepts in actual practice often requires that we first investigate how word usage correlates with the features witnessed within *unrealistically simplified but easily checked models M**.

It strikes me that the utility of Helmholtz's 'thought experiments' in geometry fit this common bootstrapping pattern rather nicely: specifically, his patently unrealistic postulates with respect to geometrical optics and rigid body movement allow a rough appraisal of 'what it's like to live in a non-Euclidean world' in a manner that makes it prima facie plausible that mathematics of this type might eventually find valid empirical application. As such, we can explain why such thought experiments 'open up possibilities' essential to conceptual advance exactly in the manner that Friedman emphasizes, without thereby demanding that our *bootstrapping models* M* must embody any specific prerequisites operative within the more realistic M that science eventually embraces. But to accept this point of view, one must adopt WS's milder position that semantics-for-the-sake-of-soundness-proofs should not be viewed as *completely capturing* in a traditionalist manner the hypothetical 'established contents' of the words under examination. Instead, the correlational studies embodied in successful soundness proofs should be primarily regarded as important vehicles for *reorienting established usage* towards higher standards of performance, exactly in the 'get me the right numbers' vein that the 'scientific philosophy' tradition has always emphasized. The conceptual utility of the classic geometrical thought experiments reflects the observation that even correlational studies with respect to somewhat unrealistic modeling families M* can serve parallel redirectional purposes.

As we first observed, DofR sounds a vital rallying call to arouse philosophy's battalions from their complacent slumbers, no longer serving as 'critics of concepts' in the great traditions of 'scientific philosophy'. Unfortunately, DofR's adoption of neo-Kantian vocabulary, if left unmollified, is apt to make this resumption of duties appear a rather daunting challenge, for present day academic philosophers, given their less-than-total immersion in the mysteries of nature, are unlikely to suggest 'framework' alterations able to assist modern physics directly on its way in the manner

[23] In the fullness of time, proper soundness results may be established for M, but their articulation often demands a prior understanding of M-structure behavior obtained from an 'unfounded' reliance upon inferential tools such as R.

of the non-Euclidean revolution. Friedman's borrowings of positivist terminology like 'the relativized a priori' make us wonder, 'Is he asking us to dream up *wrong-headed demands* upon physical theory à la Mach or Poincaré in the hopes that someday a proper physicist will be inspired to do something *good* after reading our crude philosophical follies?' Such rhetorical difficulties, it strikes me, trace to the fact that the positivists never found an adequate way to accommodate the complex 'coming at us in sections' qualities of real life conceptual advance. True, as Friedman emphasizes in his LMS note, they fully recognized that science's prevailing 'semantic prerequisites' are apt to vary from one historical epoch to another, but their heavy reliance upon *logic* as the chief organ of conceptual clarification impeded an adequate treatment of the *partial reorientations* of practice highlighted here. Thus their optimistic assumptions that able logicians would surely be able to iron the 'conceptual problems of classical mechanics' through careful axiomatization seem rather comical in retrospect, for the inferential woes trace to much deeper structural difficulties than that. (This is a major theme within WS.)

But once we no longer require that improving work upon 'conceptual contents' must operate at the 'total framework' level, then training of a traditional sort can render philosophers particularly skilled in diagnosing the types of tension that inevitably arise when one basin of usage rubs against another, as well as suggesting 'thought experiment' arrangements that might improve local strands of usage in unexpected ways. By returning to this familiar 'analytical' mode, philosophy might plausibly display the same 'critic of concepts' sagacity that one finds in great twentieth-century thinkers such as J. L. Austin and Wittgenstein (and, before them, Leibniz, Berkeley, Reid, and the 'scientific philosophers' of the late nineteenth century). Indeed, through precisely training our attention upon conceptual conflicts as they arise across a wide variety of disciplines, we might offer wary but useful advice in the wandering ways of words to specialists within their particular fields, even if we lack the empirical specifics required to unravel their localized dilemmas fully. In operating thus, philosophy can serve to 'open up possibilities' in exactly the manner Friedman suggests without requiring the extraordinary powers of intellectual engineering that the old positivist rhetoric seems to demand. Unfortunately, through the same Quinean shifts in orientation of which LMS complains, present day philosophical training is often dismissive of the traditional diagnostic skills required to resolve subtle conceptual conflicts in this manner.

Since I have just evoked Wittgenstein's name, it is worth indicating that the overall point of view suggested here is rather different than he seems to favor, judging by his evident approval of the final sentence in the Hertz quotation above. For convenience, let me cite it again:

> When these painful contradictions are removed, the question as to the nature of force will not have been answered; but our minds, no longer vexed, will cease to ask illegitimate questions. (1952: 8)

Like Hertz and myself, Wittgenstein sees language use as naturally evolving into expanding patches that eventually come into tension with one another. Like me but unlike Hertz, he does not expect that these tensions can always be 'solved' by discovering an overarching replacement theory. Yet Wittgenstein appears to think that, ultimately, all the philosopher can do is *pare back* the conflicting patches of practice so that they no longer bind against one another. He then leaves matters there, 'ceasing to ask illegitimate questions.' Ultimately, Wittgenstein thinks that there is no explanation for why these local patches exist in the first place: 'these are our practices,' etc. Such thinking exemplifies Wittgenstein's notorious 'quietist' streak; it appears to originate in that same raw suspicion of classical conceptual analysis that is expressed in the Hertz quotation.[24]

But I don't see our issues in this manner at all. We *can* eventually 'understand our practices' (although the answers often prove surprising) through suitably diagnosing the strategic complications that our awkward placement within nature necessitates. To evoke a parallelism discussed at length in WS, nineteenth-century mathematicians eventually uncovered the strategic techniques that allowed James Watt to devise an abundance of excellent improvements to the steam engine, even though the *rational basis* behind his search techniques remained completely opaque to him and his immediate followers. In providing these rationalizing underpinnings for otherwise mysterious 'practices,' we commonly rely upon the full resources of modern science and mathematics to weave a *physically convincing story* of what our linguistic circumstances are actually like. And developing unexpected correlational pictures of this ilk seems exactly the sort of thing in which philosophy should be engaged.

[24] Or if there is an 'explanation,' it is not *philosophy*'s job to provide it. I find such attitudes preposterous. For my tentative speculations on Wittgenstein's attitudes towards science, see Wilson (1997).

In fact, our 'mildly transcendental placement' in nature *demands* that we continually re-evaluate our 'practices' from a correlational point of view. If philosophy becomes unwilling to participate in such projects, well, so much the worse for its intellectual utility. Wittgenstein seems to have fallen into the same trap of one-sided thinking about 'conceptual content' as betrayed the positivists: 'Thinkers like Russell are wrong to fancy that word/world ties can be wholly forged through the direct grasp of 'content'; only linguistic practices can turn the trick.' Well, yes, the sinews of word/world connection are mainly woven through the enlarging entanglements of successful linguistic routine with worldly affairs, not baptismal intent. Nonetheless, a 'semantical point of view' upon these ties, such as provided within our correlational pictures, can prove a great boon in helping us enlarge and refine these capacities and for warning us of the applications where such methods should not be trusted. As such, a 'semantical picture' usually enjoys a *normative edge* over accepted practice, in the sense that our inferential practices must be continually hedged and curbed as a side consequence of our 'mildly transcendental' placement within nature. A correlational picture, after all, simply represents a mathematized model of what that placement is actually like. For such reasons, classical thinkers like Russell (and Tarski) were right to stress the linguistic importance of the *direct correspondences* between words and world, as a normative issue above and beyond the nature of our current practices.

On the other hand, Russell et al. erred in fancying that reliable 'semantic pictures' can be always constructed through armchair musings upon 'meanings' alone, for such correlational portraits, on the present view, should be cherished rather as valuable *instrumentalities of conceptual reorientation*, commonly in the radical manner that Friedman (and 'scientific philosophy' more generally) has identified as essential to intellectual progress. As such, novel 'semantic pictures' carry considerable normative or corrective weight vis-à-vis past practices, but they should not be viewed, pace the 'absolutist' assumptions of Dummett and others, as thereby bearing the full freight of 'meaning' for the terms they diagnose. In fact, 'meaning', as we ordinarily employ the word, should not be regarded as carrying less of a contextualized 'content' than most everyday terms of macroscopic classification, whose supportive significance commonly shifts from one occasion to another. As an ersatz totality, the full dimensions of 'meaning' remain resolutely lodged within the full array of pragmatic entanglements that successfully stitch word usage with the physical

circumstances they address. Successful correlational pictures merely add a partial and fallible *corrective adjunct* to this distributed bulk.[25]

But once our philosophical treatments of 'conceptual content' abandon their improving ambitions and we fancy that we are merely reporting upon 'what we know when we understand a language,' the components within a valid 'semantic picture' can easily seem vapid and uninformative *when applied uncritically* to one's home language, especially if one worries exclusively about the 'soundness' of logical rules of inference found there. Thereby engendered are the currently popular tropisms towards 'deflationism' with respect to 'reference' and 'truth' (including those encountered in Wittgenstein's own writings). The best curative for this undervaluing of correlational studies is to recognize the *reorientational assistance* that a well-diagnosed 'semantic picture' often provides, through implementing the hard-nosed policy of ignoring traditionalist 'loyalty to conceptual contents' demands that have always proved a hallmark of the 'scientific philosophy' movement. Like Friedman's thought experiments, such improved word-to-world constructions can provide the crucial semantic pivots that open unanticipated dimensions of exploration to the 'free creativity' of the scientist, without pretense that these useful and reorienting assemblies thereby capture 'absolutely everything essential to meaning.'

As noted above, the role of set theory and logic comes out differently on this account than according to the 'logic of concept formation' traditions traced in LMS. It is certainly true that most nineteenth-century logicians regarded the setting of proper standards for the 'construction of concepts' as comprising a more central aspect of their designated tasks than merely codifying the patterns we now call 'first order reasoning.'[26] Approached from this angle, set theory represents the desired theory of *concept formation* rendered extensional, which is exactly the manner in which the positivists conceived these issues. But our correlational approach evokes set theoretic ideas rather as a means of articulating the de facto natural relationships that ultimately render a language's descriptive lexicon useful to its employers,

[25] In the early days of Montague grammars, traditional linguists often objected, 'What possible connection can *sets* have with knowing the *meaning* of "red"?' It strikes me that a proper reply should center upon the *redirectional aspects of corrective semantics* as surveyed here. *Formal semantical description* should be accorded a significant role within our treatment of language, without pretending that it fully answers to 'meaning' in all of its expected dimensions.

[26] Besides Friedman's own writings, useful surveys of this background can be found in Carus (2007: chs. 2–3) and Wolfgang Carl (1994: chs. 1–2).

quite independently of whether they have any just inkling of how those strategic correlations actually unfold. (As we've seen, we commonly picture language's workings wrongly.) Following the lead of the applied mathematicians, we should no longer concern ourselves with establishing an *architectural toolkit* that any would-be system builder can exploit in framing concepts to suit the 'free creativity' demands of science. From the present point of view, it is rather dubious that such a foundationalist project is feasible at all, for devising linguistic strategies that can successfully function within a complex and uncooperative natural environment demands continual syntactic experimentation and mathematical study, whose convoluted contours are very hard to anticipate a priori. To capture the myriad coordinative manners in language can effectively entwine itself with worldly events, we require the basic notions of set theory—*mapping* and *limit*—but not because we trying to supply a general theory of the 'internal contents of concepts' in a traditional manner. Our point of view is resolutely *externalist*: we attempt to assess our de facto computational position in nature in the hope of devising more sophisticated stratagems that might work more effectively. So we must learn how our reasonings and computations currently *map* to the world they address, in the general fashion in which we examined Euler's rule. And because this placement often proves of a 'mildly transcendental' character, the notion of *limit* also becomes vital as well, for our Eulerian computations $\mathbf{C}^{\mathbf{E}}$'s only approach their target **C**'s in that asymptotic fashion. Indeed, a look at the historical record will find set theoretical thinking gradually seeping into mathematical analysis in the middle nineteenth century for precisely these reasons, which are quite detachable from the 'logic of concepts' motivations recounted in LMS.[27]

While on this topic, it is worth entering a quick complaint with respect to some widespread misapprehensions. Many self-styled 'physicalists' presume that they adequately understand the 'physical inventory' of the world and that no sets or allied 'abstract objects' lie among them. (They merely represent 'mathematical artifacts,' whatever those might be.) Something has plainly gone amiss in such thinking. Surely, the manner in which an Euler-style computation relates, or fails to relate, to the reality it addresses comprises a vital empirical characteristic of the world in which

[27] As I understand her, Penelope Maddy (2007) regards such attitudes as paradigmatic of what she calls 'second philosophizing.'

we live. And the notions of *map* and *limit* represent the natural vocabulary in which such relationships should be discussed. It cannot represent a plausible requirement upon a reasonable 'physicalism' that it should relinquish the very terminology ('infinite computational set serving as limiting envelope') one employs to register garden variety forms of computation-to-world relationship.[28]

In sum, if we no longer demand that science advance in great blocks of coherent framework built upon well-articulated hunks of theory-to-measurement presupposition, we can better respect the fact that real life science only 'comes at us in sections' while underwriting Friedman's basic claim that philosophy should resume its former role as 'critic of concepts,' in the best traditions of scientific philosophy. As suggested above, some of DofR's terminological borrowings from the positivist heritage (e.g. 'relativized a priori') strike me as less than ideal, for fleshing out an accurate portrait of our 'computational position in nature' suggests a chastened scientific realism to me, rather than the modernized Kantianism that Friedman espouses.[29] But these divergencies may prove more terminological in nature than substantive. However that may be, Friedman and I fully concur that meditating exclusively upon DOG and DOORKNOB as such traits appear in undemanding domesticity will rarely led the modern student to a proper appreciation of the *difficult practical dilemmas of conceptual guidance* that have always animated the 'scientific philosophy' movement. Such a complacent myopia, we think, is unwise; sound philosophizing requires greater critical grit beneath its wheels than mere DOG and DOORKNOB encourage.

[28] Nor should one wish to code such relationships in unnatural ways, à la Hartry Field's *Science Without Numbers* (1980). In addition, it should be observed that *understanding* how these mappings operate strategically usually requires that we embed these maps within a richer mathematical setting. For example, Cauchy's original questions on series convergence require that we study their behavior upon the complex plane (indeed, upon Riemann surfaces) as well. Throughout his career, Robert Batterman has stressed the fact that 'understanding' in science commonly requires the interpolating intervention of such 'abstract' mathematical structures (cf. his *The Devil in the Details*, 2001).

[29] To a realist *au fond* such as myself, it is better to retain a conception of 'objectivity' as 'correlates successfully with the world' rather than embracing the ersatz 'objectivity as shared human standards' that Kantians usually substitute in its stead. (See DofR, p. 67, for an expression of the latter inclination.)

References

Baker G. P. and Hacker, P. M. S. (eds.) (2005). *Wittgenstein: Understanding and Meaning*. Oxford: Wiley-Blackwell.

Batterman, R. (2001). *The Devil in the Details*. Oxford: Oxford University Press.

Browder, F. (ed.) (1983). *Mathematical Developments Arising from Hilbert Problems*. Providence: AMS Press.

Carl, W. (1994). *Frege's Theory of Sense and Reference*. Cambridge: Cambridge University Press.

Carus, A. W. (2007). *Carnap and Twentieth-Century Thought*. Cambridge: Cambridge University Press.

Dummett, M. (1978). The justification of deduction. In *Logic and Other Enigmas*. Cambridge, MA: Harvard University Press.

Clifford, W. K. (1879). *Lectures and Essays*. London: MacMillan.

Field, H. (1980). *Science Without Numbers*. Princeton: Princeton University Press.

Friedman, M. (2001). *Dynamics of Reason*. Stanford: CLSI Publications.

—— (2010a). Logic, mathematical science and twentieth century philosophy. *Noûs,* 44 (3): 520–44.

—— (2010b). Replies. In M. Dickson and M. Domski, (eds.) *Discourse on a New Method*. La Salle: Open Court.

Helmholtz, H. (1977). On the origin and significance of the axioms of geometry. In R. S. Cohen and Y. Elkana, (eds.) *Hermann von Helmholtz: Epistemological Writings*. Dordrecht: D. Reidel.

Hertz, H. (1952). *The Principles of Mechanics*, translated by D. E. Jones and J. T. Walley. New York: Dover.

Jammer, M. (1962). *The Concept of Force*. New York: Harper's.

Lutzen, J. (2005). *Mechanistic Images in Geometrical Form*. New York: Oxford University Press.

Maddy, P. (2007). *Second Philosophy*. Oxford: Oxford University Press.

Manin, Y. (1981). *Mathematics and Physics*. Boston: Birkhauser.

Reid, C. (1991). Hans Lewy. 1904–1988. In P. Hilton, F. Hirzebruch, and R. Remmert, (eds.) *Miscellanea Mathematica*. Berlin: Springer-Verlag.

Wagner, P. (ed.) (2011). *Carnap's Ideal of Explication and Naturalism*. London: Palgrave Macmillan.

Wilson, M. (1997). Wittgenstein: *Physica sunt, non leguntur*. *Philosophical Topics*, 25: 289–316.

—— (2006). *Wandering Significance*. Oxford: Oxford University Press.

—— (2010). Frege's mathematical setting. In M. Potter and T. Ricketts (eds.) *The Cambridge Companion to Frege*. Cambridge: Cambridge University Press.

—— (2011). Of whales and pendulums: A reply to Brandom. *Philosophy and Phenomenological Research*, 82 (1): 202–11.

8

Neo-Kantianism, Scientific Realism, and Modern Physics

Michael Friedman

The Kantian conception of the rationality and objectivity of human cognition, as I understand it, is a reaction (among other things) to the spectacular success of Newton's *Principia*. In particular, Kant's conception of our two forms of sensible intuition—space and time—and our pure intellectual categories of substance, causality, and community is modeled on the Newtonian concepts of absolute space and time, on the one side, and the Newtonian concepts of mass, force, and mutual interaction, on the other. Newton's theory of universal gravitation, as applied to the observable relative motions of the bodies in the solar system, provides Kant with a model for his own conception of a system of (phenomenal) substances in space and time interacting in accordance with the categories. Space, time, and the categories, as universal a priori structures of the human mind, thereby provide a universal a priori framework for our cognition of nature—at all times and in all places—and thus secure the trans-historical rationality and objectivity of all possible sciences developed by us.

It is perhaps clear and understandable why Kant took the Euclidean geometry of space to have such a 'transcendental'—knowledge-constituting—status. But why did he accord essentially the same status to the Newtonian concepts of mass, force, and mutual interaction—which, from Kant's point of view, instantiate the fundamental a priori categories of substance, causality, and community? Here, on my telling, Kant is responding to the problem of absolute space, time, and motion as it arises in the context of Newton's original theory. Rather than taking absolute space, time, and motion to be already well defined (by the explanations in

Newton's famous Scholium to the Definitions) prior to the Newtonian laws of motion governing mass, force, and mutual interaction, Kant views these laws as rather defining a privileged frame of reference (based on the center of mass of the solar system) within which such motions are themselves first defined. Kant thereby views the Newtonian laws of motion as what we would now call an *implicit definition* of the privileged space, privileged time, and privileged state of (ideal force-free) motion relative to which Newton's theory is true.[1] Indeed, the modern concept of *inertial* reference frame, developed in the late nineteenth century as a solution to the problem of Newtonian absolute space, time, and motion, proceeds by just such an implicit definition: the inertial frames, in this tradition, are just those in which force-free bodies follow straight lines with constant velocity, and in which every (true) change of momentum representing a deviation from this state is precisely counterbalanced by an equal and opposite change of momentum in the source of the force producing the change. I believe that Ernst Mach adopted a version of Kant's way of formulating this solution in the late nineteenth century and thereby connected Kant's original solution with the modern concept of inertial frame.[2]

Also in the late nineteenth century, however, this classical concept of inertial frame came under increasing pressure due to developments from outside the provenance of classical Newtonian theory: namely, the development of Maxwellian electrodynamics. For it now appeared that the classical principle of Galilean relativity—already well known to Newton as Corollary V to the laws of motion in the *Principia*—might not be correct. In this context, it appeared, the inertial frames defined from within the Newtonian tradition are not all equivalent after all, and there should be a privileged rest frame relative to which the velocity of light (and all other electromagnetic disturbances) is exactly *c*. Of course, repeated attempts experimentally to distinguish different inertial frames from one another on this basis all failed, and the solution, from a modern point of view, was finally given by the creation of Einstein's special theory of relativity in 1905. Moreover, Hermann Minkowski then gave a geometrical presentation of this theory within the tradition of Hilbertian axiomatics, wherein

[1] For details of this way of understanding the relationship between Kant and Newton see Friedman (1992). See also Friedman (2004).

[2] For the development of the modern concept of inertial frame see DiSalle (2006). For its relation to the work of Mach, in particular, see DiSalle (2002).

the Einsteinian electrodynamic framework for space, time, and motion appears as a (four-dimensional) geometrical alternative to the parallel space-time geometry for classical Newtonian theory implicitly defined by the Newtonian laws of motion.[3]

In this new, Einsteinian-Minkowskian framework, the fundamental relation between spatially separated events of (distant) absolute simultaneity is now relativized to a choice of inertial reference frame (as now implicitly defined by Einstein rather than Newton). But this relation, in turn, was fundamental to Newtonian gravitation theory, since gravitational interaction, on this theory, takes place immediately—that is, simultaneously—across arbitrarily large spatial distances. Einstein therefore needed radically to revise gravitation theory as well. And he did this, in the years 1907–15, by taking the truly revolutionary step of describing gravitational interaction by a version of (four-dimensional) Minkowski geometry in which space-time is curved rather than flat—a variably curved non-Euclidean four-dimensional geometry in which the curvature depends on the distribution of mass-energy and 'freely falling' bodies (affected by no forces other than gravity) follow four-dimensional geodesics of this geometry. Freely falling trajectories, in accordance with what Einstein called the principle of equivalence, thereby replace the inertial trajectories of Newtonian theory, and inertia itself becomes a geometrical phenomenon rather than simply a mechanical one.

These developments involving both special and general relativity pose a deep challenge to Kant's original conception of scientific objectivity. And they were explicitly considered as such in works of both neo-Kantian and logical empiricist philosophers soon after the creation of the general theory—works by Moritz Schlick, Hans Reichenbach, Rudolf Carnap, and Ernst Cassirer. At issue was whether the original Kantian conception should simply be discarded (as argued by Schlick) or rather generalized and revised (as argued, in one form or another, by the others). One alternative, suggested by Cassirer (1921/23) and explicitly developed by the early Reichenbach (1920/65), was to *relativize* the Kantian idea of a priori 'constitution' and a priori constitutive principles, so that they vary from theory to theory as science changes and develops but still play the characteristically Kantian role of *defining* the space of real (mathematical and physical) possibilities within which properly empirical laws can then be

[3] For details of Minkowski's work and its relation to the Hilbertian axiomatic tradition see Corry (1997).

formulated.[4] My *Dynamics of Reason* (2001) picks up and develops this alternative. The idea is that, just as Newton's three laws of motion made the theory of universal gravitation (and thus the law of universal gravitation) possible as a properly empirical statement (by implicitly characterizing the spatiotemporal structure relative to which gravitational changes of momentum are first well-defined), something similar happens in the theory of relativity. Einstein's 'kinematical' definition of simultaneity (appealing to the constancy and invariance of the velocity of light) implicitly characterizes the relativistic (Lorentzian) inertial frames within which Maxwell's equations are now true. And Einstein's principle of equivalence characterizes the privileged state of 'freely falling' motion in terms of which the mathematical geodesics of the variably curved space-time geometry of the general theory first acquire an empirical or physical significance.

The result is a relativized Kantian-style conception of scientific rationality or objectivity that develops and changes with the progress of empirical science. But this, as I have also argued, is very similar to the relativization of objectivity to 'paradigms' during Kuhnian scientific revolutions. The transition from Newton to Einstein is Kuhn's main example of what he takes to be the conceptual incommensurability characteristic of scientific revolutions.[5] And Kuhn's historiographical approach has clear roots in the neo-Kantian tradition—in the work of such thinkers as Emile Meyerson, Alexandre Koyré, and (once again) Ernst Cassirer.[6] It is no surprise, therefore, that Kuhn took to characterizing his conception as 'Kantianism with movable categories' and, very late in his career, also recognized its kinship with the relativized a priori found in early logical empiricism.[7] On this basis, I have reinterpreted the problem of Kuhnian conceptual incommensurability in more Kantian terms. The problem is that Einsteinian relativity articulates a radically different space of conceptual possibilities (both mathematical and physical) that were simply *impossible* relative to Newtonian theory. In particular, whereas one can indeed represent the structure of Newtonian gravitation theory as an approximate special case of Einstein's theories (by, for example, letting the speed of light

[4] I discuss Reichenbach (1920/65) and Cassirer (1921/3), in relation to Schlick, Carnap, and others, in my *Reconsidering Logical Positivism* (Friedman, 1999).

[5] See Kuhn (1963), Chapter IX, 'The nature and necessity of scientific revolutions'.

[6] See Friedman (2003).

[7] See Kuhn (1993).

go to infinity and considering the solar system as an isolated system not affected by anything outside it), the possibility-defining constitutive principles of Newtonian theory (the laws of motion) come out as mere empirical conditions holding in a parochial approximate limit: we can thereby uniquely define the Newtonian inertial frames and gravitational potential, but only relative to a now entirely contingent distribution of matter.

The problem then becomes explaining how the new space of conceptual possibilities can rationally evolve from the old one. And I have addressed this problem, in turn, by an historical narrative depicting the various and intricate ways in which the mathematical and physical developments leading from Newton to Einstein were inextricably entangled with a parallel set of developments within scientific philosophy principally involving Ernst Mach, Hermann von Helmholtz, and Henri Poincaré—all of whom, in one way or another, continue the original Kantian legacy in a revised form. The upshot of this narrative is that a cutting-edge turn-of-the-century classical physicist—such as Henri Poincaré—had very good reasons *in his own terms* for taking Einstein's expansion of the classical space of conceptual possibilities seriously: for accepting it, that is, as itself really possible.[8] This is how, beginning with Kant and his relation to Newton, I have been first led to the importance of Hilbert-style implicit definitions in explaining how our fundamental mathematical concepts of space, time, and motion acquire their empirical meaning, then to the relativized a priori and its connection with Kuhn's theory of scientific revolutions, and finally to a solution of the (reinterpreted) problem of conceptual incommensurability emphasizing the mediating role of scientific philosophy.

The project of Mark Wilson's *Wandering Significance* (2006) has much in common with that of my *Dynamics of Reason* (2001). We both address problems of interest to contemporary philosophy of science—and to contemporary philosophy more generally—by presenting detailed accounts of particular episodes from the history of the mathematical sciences. Moreover, we both concentrate on events at the end of the nineteenth and beginning of the twentieth centuries where key developments in modern mathematical physics were intimately entangled with late nineteenth-century 'scientific philosophy'—involving such figures as Hermann von Helmholtz, Ernst Mach, James Clerk Maxwell, Henri Poincaré, Heinrich

[8] For the most detailed and up-to-date version of this historical narrative see Friedman (2010).

Hertz, and Pierre Duhem. These developments had a defining impact, in turn, on the development of early twentieth-century scientific philosophy in the work of Bertrand Russell, Moritz Schlick, Rudolf Carnap, and many others. Wilson and I emphatically agree that both contemporary philosophy of science and contemporary philosophy more generally have much to learn from revisiting these earlier developments.

But there are important differences between our two projects when it comes to evaluating the broader metaphysical and epistemological significance of the scientific developments on which we focus. I place my historical narrative, as explained, within an explicitly Kantian philosophical tradition. In doing so I operate with a conception of scientific objectivity and rationality emphasizing intersubjective agreement among rational human knowers rather than a correspondence between our scientific representations and a mind-independent external world:

> For me, the main problem posed by post-Kantian scientific developments is precisely a challenge to the idea of a universal (trans-historical) rationality, paradigmatically expressed in modern mathematical science, which can serve as a basis for rational argument and communication between all human beings. . . . Just as Kant's original defense of scientific rationality did not proceed on the basis of what he called 'transcendental realism,' the idea that our human system of representations somehow corresponds to an otherwise entirely independent realm of 'things in themselves,' the present defense does not proceed on the basis of 'scientific realism.' (Friedman, 2001: 117–18)

Wilson, by contrast, explicitly distances himself from such a Kantian (or neo-Kantian) perspective in endorsing scientific realism:

> Throughout this book, I take the facts of mathematics pretty much for granted. However, the notion that this subject must assume the role of *regulative principle* prior to any description of the world in physical terms represents a vital aspect of neo-Kantian tradition, as aptly emphasized by my friend Michael Friedman. In this book I have not attempted to dabble in topics so grand as these; I have instead considered concepts entirely from a scientific realist point of view. I do believe that the easy road to neo-Kantianism has been paved, historically at least, by strong reliance on veil of predication related claims. (Wilson, 2006: 84)

In the same vein, Wilson considers (and rejects) contemporary attempts to supply us with ersatz notions of 'objectivity' derived from Kant: '[T]hese surrogate proposals follow Kant in claiming that a defensible notion of conceptual objectivity should turn upon our abilities to reach a classificatory or truth-evaluative accord with our fellow men: proper "objectivity" in

classification represents a matter of *inter-personal* agreements rather than correspondence to unsullied data' (Wilson, 2006: 80).[9] As Wilson rightly emphasizes, this rejection of correspondence on behalf of intersubjectivity is characteristic of the early twentieth-century neo-Kantian scientific philosophy developed by Cassirer: '[Cassirer's] approach presumes a rejection of straightforward realism with respect to either the physical world or mathematics, following the usual neo-Kantian inclination to treat scientific objectivity as the sharing of investigative standards between different public parties, rather than direct correspondence with empirical reality. Allied themes remain popular in philosophical circles today, although I will have no truck with them myself' (Wilson, 2006: 153).[10] Wilson's friendly distancing of himself from the project of my *Dynamics of Reason* therefore goes hand in hand with his respectful yet firm rejection of Cassirer's neo-Kantian approach.

A related important difference between our projects concerns the central place I give to Kuhn's theory of scientific revolutions—which I take very seriously indeed as constituting the main threat to scientific rationality and objectivity I want to overcome. In particular, whereas I want to acknowledge the importance of Kuhn's notion of conceptual incommensurability, this notion represents the final degeneration of the approach to scientific concepts represented by the Hilbertian conception of 'implicit definitions' within a global axiomatic scientific theory according to Wilson. This conception, on his account, leads ultimately to Quine's 'hazy holism,' and Kuhn's picture of an unresolvable clash between conceptually incommensurable scientific paradigms is its logical outcome: '[My] discussion will display a mesoscopic emphasis that falls between the attention to individual word meaning typical of the classical tradition and the sprawling webs of belief favored by Quine and his cohorts. In my diagnosis, it is the intimations of intensionality arising in the middle range that most commonly occasion the familiar puzzlements of *ur*-philosophy, as well as inducing the scientific impasses that Kuhn mistakenly characterizes as the clash of paradigm-addled mind sets' (Wilson, 2006: 306).[11] Wilson concretely illustrates this point by

[9] Wilson is here considering contemporary work by Gary Ebbs and Crispin Wright.

[10] It appears that the 'allied' contemporary philosophy to which Wilson refers includes my own work, as well as that of Ebbs and Wright noted above.

[11] The connection between Kuhn's view, Quine's 'hazy holism,' and Hilbertian axiomatics emerges on pp. 157–77, 279–86.

considering the turn-of-the-century clash between Pierre Duhem and Lord Kelvin concerning the role of classical-mechanical molecular theory in accounting for the observed behavior of macroscopic matter: '[A]s modern academic philosophers and historians we can properly fault *ourselves* for having not profited better from classical physics' travails by continuing to see such disputes as battles of paradigms and webs of belief, rather than puzzled reactions that arise when unsuspecting practitioners confront the delicate filigree of patchwork arrangements typical of successful classes of reduced variables' (Wilson, 2006: 368).[12]

I shall return below to exactly what Wilson means by such a 'delicate filigree of patchwork arrangements,' and to why he takes this idea both to support scientific realism and to be tension with the Hilbertian emphasis on implicit definitions. But I first want to say something about the origin and motivations of Wilson's project in order to facilitate its comparison with mine. I hope to show that the differences between us concerning neo-Kantianism, scientific realism, implicit definitions, and Kuhnian paradigms are not as irreconcilable as it might first appear. I believe, in the end, that our two projects are complementary rather than incompatible—focussing on different but equally important aspects of the historical development of modern physics from Newton to the early twentieth century.

Wilson's philosophical journey began when he was a student of Hilary Putnam's at Harvard in the late 1960s and early 1970s. He was attracted to the 'new theory of reference' articulated by Putnam and Saul Kripke, and by the robust scientific realism within which Putnam (at that time) placed this theory. Wilson was convinced early on, however, that this new theory of reference was much too simple-minded in such pronouncements as '"water" refers to H_20,' for example, and that, more generally, our ordinary terms of macroscopic conceptual classification relate to the physical world in a much more complicated fashion. Wilson's voracious reading in the history of physics and mathematics, engineering, the theory of differential equations, and virtually every other subject concerned with how the physical world and our linguistic representations of it interact with one another, then propelled him on an incredibly detailed and breathtakingly original investigation of these matters. The first fruit was

[12] The connection with Hilbertian axiomatics and its later degeneration at the hands of Quine and Kuhn emerges on pp. 193–203, to which I shall return below.

his great paper, 'Predicate Meets Property,' appearing in 1982. The latest is *Wandering Significance*.

In the latter, Wilson sums up his relationship with Putnam as follows:

> In these respects, my quasi-biomechanical recipes for unraveling the intensional characteristics of predicates are distinctly 'externalist' or 'naturalist' in flavor (although I do not care for either of these popular phrases much). An allied externalist orientation seems evident in the 1974 Putnam essay ['The Meaning of "Meaning"']... Indeed, I was a student of Putnam's in the relevant period and many of my musings can be fairly credited to (or blamed upon) the vital spark of anti-classicism that I derived from his teachings, as well as the mode of straightforward scientific realism that his essays of the same period seemed to embrace (he has subsequently denied that this realist stance represented his fully considered point of view). Unlike the Putnam of 1974, however, I do not embrace the supplementary mechanisms of original intention (e.g. 'I hereby baptize this liquid, whatever it is, as "water"') that Putnam includes in order to insure that predicates such as 'is water' maintain invariant extensions over their extended careers ... I reject these doctrines because they seem descriptively inaccurate and inconsistent with fundamental tenets of a reasonable anti-classicism ('liquid', after all, behaves even more irregularly in its predicative fixity than 'water'). In any case, the supportive fabric of facade I shall defend displays rather different characteristics than any scheme that Putnam contemplates. (Wilson, 2006: 136)[13]

In short, Wilson accepts Putnam's early scientific realism and related project of delineating the actual word-world correlations exhibited in our linguistic behavior (especially our scientific linguistic behavior), but he replaces the simple-minded correlational mechanisms envisioned by Kripke and Putnam by what he calls 'the supportive fabric of facade.'

Wilson focusses on predicates of macroscopic physical classification—concepts such as *force*, *hardness*, *rigidity*, and *redness*—which have been taken to be problematic in various ways since the birth of classical physics in the seventeenth century. Continuous macroscopic bodies (as opposed to the point-particles or points-masses of traditional Newtonian mechanics) became subject to mathematical-physical treatment during the eighteenth and nineteenth centuries, principally through the mathematical theory of partial differential equations. This theory then led to some of the most beautiful and important developments in nineteenth-century mathematics (Fourier series, complex analysis, Riemann surfaces, Sturm-Liouville existence and uniqueness theory), which, in turn, eventually led to the development of

[13] Wilson is referring to Putnam's classic paper, 'The Meaning of "Meaning"' (1974).

modern set theory as the overarching framework within which solutions to such equations can be rigorously established and investigated. And, at the same time, the evolving mathematical theory of partial differential equations played an essential role in the description of macroscopic matter in classical continuum mechanics, where the essentials of Newton's conceptual apparatus were extended to this new realm. The Newtonian conceptual apparatus (the concept of *force* in particular) was also essentially extended, in different yet related ways, in the investigation of such 'constraints' as *rigidity* by (among others) Heinrich Hertz.

Profound conceptual problems accompanied these extensions and led to puzzling shifts and anomalies in the behavior of the relevant classical concepts, especially when applied to macroscopic matter. The upshot, for Wilson, is that the concepts of classical Newtonian mechanics do not form a self-sufficient system when applied to this domain: aside from the traditional mechanical concepts (force, mass, and so on), alien concepts from outside mechanics proper (concepts belonging to chemistry and thermodynamics, for example) proved to be absolutely indispensable as well. The classical description of macroscopic matter thereby showed itself to be no self-sufficient and self-consistent *theory* at all, but rather what Wilson very helpfully calls a *theory facade*: a makeshift patchwork of mathematical descriptive devices sitting on top of the true (as far as we now know) microscopic reality depicted by quantum mechanics, much as the true spherical shape of the earth can be described by a patchwork (or atlas) of flat two-dimensional maps related to this true spherical shape by a variety of different mathematical projections. And it is precisely this situation that underlies Wilson's anti-Kuhnian (and pro-'scientific realist') diagnosis of the dispute between Duhem and Kelvin remarked upon above.[14]

We are now in a position to understand more precisely why Wilson has difficulties with the Hilbertian notion of implicit definition, particularly when taken as a complete explanation of the meanings of the fundamental concepts of mathematical physics. Indeed, as Wilson points out, Hilbert included a successful axiomatization of classical mechanics as a desideratum in his famous list of turn-of-the-century mathematical problems, and what we now see, on Wilson's telling, is that no such self-sufficient axiomatization is possible. Classical mechanics is simply a collection of local patches of

[14] See again n. 12 above, the significance of which will become clearer immediately below.

mathematical representation constituting an atlas-like theory facade: it is not something that could possibly be successfully organized as a global axiomatic theory.[15] In this case, the whole idea of a complete system of implicit definitions delineating a Kuhnian paradigm or conceptual framework makes no sense. In particular, there is nothing that characterizes what we might be tempted to call the full space of possibilities allowed by classical mechanics—the set of classical-mechanically possible worlds. It appears, therefore, that, if we truly want to understand the conceptual pressures leading to the breakdown of classical physics at the turn of the nineteenth century, we should abandon Kuhn—along with all other neo-Kantian projects—and embrace the robust scientific realism underlying Wilson's detailed and delicately filigreed accounts of actual word-world correlations.

My response to this last suggestion is that the conceptual pressures emphasized in Wilson's account—those having to do with the relationship between classical-mechanical molecular theory and the behavior of macroscopic matter—are not the only factors leading to the breakdown of classical physics at the turn of the century. And my project, more specifically, focusses on a rather different and orthogonal set of developments concerning the relationship between our fundamental mechanical concepts of space, time, and motion and the empirical phenomena they aim to describe—namely, first and foremost, the observable purely relative motions we actually experience. In particular, the problem engendered by the Newtonian concepts of absolute space, time, and motion did not concern the relationship between macroscopic matter and its microscopic constitution. The problem was rather that absolute space, time, and motion, as Newton characterizes them in the Scholium, appear not to be part of the empirical or phenomenal world at all: they have no possible causal relation to our senses. How, then, can these concepts—especially as employed in the theory of universal gravitation developed in Book III of the *Principia*—be as empirically successful as they in fact are? The Kantian

[15] Wilson is here discussing Hilbert's Sixth Problem. Its connection with Wilson's concern with 'theory facades' and his perspective on the Duhem-Kelvin dispute becomes explicit on p. 358 (2006), where Wilson explains that 'Duhem is simply insisting upon the same intelligible order that Hilbert requested in his sixth problem on mechanics discussed [on pp. 166–71].' As a culmination of the latter discussion Wilson explains (2006: 195) that 'this consideration supplies the true reason why Hilbert's sixth problem on the foundations of classical mechanics was never fully resolvable in its originally intended terms.'

answer, as later elaborated by Mach and the late nineteenth-century conception of inertial reference frame, is that the Newtonian laws of motion implicitly characterize a privileged frame of reference (fixed, for example, at the center of mass of the solar system) within which the true or absolute motions described by universal gravitation are first well defined.

In this specific context, therefore, the notion of implicit definition leads to a genuine advance in conceptual clarification. And it is important to note that this is not simply a Hilbertian implicit definition of an abstract mathematical structure (as in Hilbert's own axiomatization of Euclidean geometry, for example). Rather, it shows how an abstract mathematical structure (what we now call the structure of Newtonian space-time) actually corresponds to something in the empirical or phenomenal world. It does this by including not only pure mathematical concepts (such as Hilbert's points, lines, and planes) in the implicit definition in question, but also fundamental mechanical quantities such as mass, momentum, and force. Newton's argument for universal gravitation in Book III then shows us how to measure these quantities empirically so as actually to determine the relevant center of motion: we empirically compare the masses of the primary bodies in the solar system, for example, so as empirically to determine its center of mass. This argument *presupposes* that both Euclidean geometry and the Newtonian laws of motion are already in place, and it is in precisely this sense, in my view, that both Euclidean geometry and the laws of motion are constitutively a priori in Newton's theory.

This kind of correspondence between mathematical concepts and empirical phenomena is quite different from the word-world correlations described by Wilson's atlas-like facade structures. It is not a matter of establishing local correlations between different patches of a theory facade and aspects of microscopic physical reality, but of implicitly characterizing a global coordination between abstract spatiotemporal representations and concrete observable phenomena via what Reichenbach (and others) called coordinating principles. We do this, in the case of Newton's original theory, via his laws of motion; or, in the case of special relativity, via what Einstein called the light principle and the principle of relativity; or, in the case of general relativity, via an axiomatic characterization involving both light propagation and the behavior of freely falling bodies.[16] The

[16] Compare, in this connection, the modern axiomatic treatment (in the tradition of Hilbert and Hermann Weyl) of Ehlers, Pirani, and Schild (1972).

problems addressed by these kinds of axiomatic characterizations of spatiotemporal structure via coordinating principles are fundamentally different from those addressed by Wilson's atlas-like explanations of the patchwork behavior of macroscopic physical concepts.

In truth, the great conceptual revolutions leading from classical to twentieth-century mathematical physics, through the extraordinarily rich intervening developments in both mathematics and physics of the nineteenth century, involved both kinds of problems: those involving the most general spatiotemporal form of physical theory (general mechanics) and those involving the particular kinds of matter interacting within this form. Physics, both classical and post-classical, is essentially concerned with both kinds of problems, and a so far undeveloped area in history and philosophy science would involve carefully studying the complex interconnections between these two sets of problems. The Newtonian theory of universal gravitation applied to the solar system, for example, can proceed largely independently of the details of matter theory. Once we know the masses of the sun, the planets, and their satellites, together with the gravitational forces acting between them, the problem is essentially solved—and this is why we can present an idealized version of the solar system as a system of point-masses. (In treating the shape of the earth, however, or in studying tidal friction, this idealization is no longer adequate, and we do need to take seriously both matter theory and forces other than gravitation.) By contrast, as Wilson explains, in applying the same Newtonian laws of motion to collisions, very difficult problems in matter theory become unavoidable. For Newton describes the different outcomes of collisions between hard, elastic, and soft bodies (all of which satisfy the conservation of momentum) only by introducing a 'constant of restitution' determined purely empirically. And a genuine theoretical description of what goes on in such cases, as Wilson also explains, quickly entangles us in one of the most complicated facade-structures imaginable.[17]

Now the revolutionary transition from Newton to Einstein also essentially involved matter theory, as well as the first theoretical articulation of a fundamental force of nature other than gravity. For the problems addressed by Einstein's 'electrodynamics of moving bodies' arose squarely within Henrick Lorentz's theory of the electron, explicitly developed as an

[17] See Wilson (2006: 177–83).

answer to the question how matter and the aether (i.e. the electromagnetic field) are related within Maxwell's theory. In precisely this context, moreover, we faced very difficult problems concerning whether the electron was to be conceived as a rigid body, an elastic deformable body, or a point source—problems that are naturally grist for Wilson's mill. Nevertheless, Einstein's own approach to electrodynamics involved deliberately bracketing these problems in matter theory, at least temporarily. He aimed to construct a 'principle theory' rather than a 'constitutive theory'—one that would describe the new spatiotemporal structure required without making premature commitments in microphysics. The idea was to determine the new spatiotemporal symmetries independently of the rapidly developing details of microphysics, and then to use these symmetries to constrain this development. Einstein followed a similar approach in the creation of general relativity. In this case, as he explains in his celebrated paper 'Geometry and Experience' in 1921, he needed to arrive at a geometrical interpretation of the gravitation field, and he successfully arrived at this interpretation by applying the principle of equivalence to a uniformly rotating (and thus non-inertial) frame of reference in the context of special relativity. He therefore needed, in particular, to follow Helmholtz (rather than Poincaré) in taking the geometry of physical space to be directly reflected in the behavior of 'practically rigid bodies,' entirely independently of any theoretical consideration of their actual microscopic constitution. Thus, once again, space-time physics, on Einstein's view, must precede, and then constrain, the development of microphysics.[18]

At the time, as we know, Einstein was preoccupied with his own attempts at constructing a unified field theory, which he increasingly saw as the only acceptable alternative to the emerging theory of quantum mechanics (and therefore as the only acceptable alternative period). Twentieth-century physics, as we also know, has followed a quite different path. But the conceptual relationships between relativity (both special and general) and quantum theory (including quantum field theory) are still very much in flux. And one way to come to grips with this situation, I suggest, is through a careful study of the evolving interrelationships between space-time physics and matter theory throughout the modern period. Here,

[18] For details see Friedman (2002).

I believe, there is more than enough room for both the Kantian themes that I have been emphasizing and the extremely sophisticated elaboration of an (early-)Putnam-style scientific realism by my friend Mark Wilson.

References

Cassirer, E. (1921/1923). *Zur Einsteinschen Relativitätstheorie: Erkenntnistheoretische Betrachtungen*. Berlin: Bruno Cassirer. Translated as *Einstein's Theory of Relativity*. Chicago: Open Court.

Corry, L. (1997). Hermann Minkowski and the postulate of relativity. *Archive for the History of the Exact Sciences*, 51: 273–314.

DiSalle, R. (2002). Reconsidering Ernst Mach on space, time, and motion. In D. Malament (ed.) *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, pp. 167–91. Chicago: Open Court.

—— (2006). *Understanding Space-Time: The Philosophical Development of Physics from Newton to Einstein*. Cambridge: Cambridge University Press.

Ehlers, J., Pirani, F., and Schild, A. (1972). The geometry of free fall and light propagation. In L. O'Raifeartaigh (ed.) *General Relativity: Papers in Honour of J. L. Synge*, pp. 63–84. Oxford: Oxford University Press.

Friedman, M. (1992). *Kant and the Exact Sciences*. Cambridge, MA: Harvard University Press.

—— (1999). *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.

—— (2001). *Dynamics of Reason: The 1999 Kant Lectures at Stanford University*. Stanford: CSLI.

—— (2002). Geometry as a branch of Physics: Background and context for Einstein's 'Geometry and experience'. In D. Malament (ed.) *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, pp. 19–229. Chicago: Open Court.

—— (2003). Kuhn and logical empiricism. In T. Nickles (ed.) *Thomas Kuhn*, pp. 19–44. Cambridge: Cambridge University Press.

—— (2004). *Immanuel Kant: Metaphysical Foundations of Natural Science*. Cambridge: Cambridge University Press.

—— (2010). Synthetic history reconsidered. In M. Domski and M. Dickson (eds.) *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science*, pp. 571–813. Chicago: Open Court.

Kuhn, T. (1963). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

—— (1993). Afterwords. In P. Horwich (ed.) *World Changes: Thomas Kuhn and the Nature of Science*, pp. 311–41. Cambridge, MA: MIT Press.

Putnam, H. (1974). The meaning of "meaning". In K. Gunderson (ed.) *Language, Mind and Knowledge* (Minnesota Studies in the Philosophy of Science, vii, pp. 131–93). Minneapolis: University of Minnesota Press. Reprinted in H. Putnam, (1975) *Mind, Language and Reality* (Philosophical Papers, ii, pp. 215–71). Cambridge: Cambridge University Press.

Reichenbach, H. (1920/1965). *Relativitätstheorie und Erkenntnis Apriori*. Berlin: Springer. Translated as *The Theory of Relativity and A Priori Knowledge*. Berkeley and Los Angeles: University of California Press.

Wilson, M. (1982). Predicate meets property. *The Philosophical Review,* 91: 549–89.

—— (2006). *Wandering Significance: An Essay on Conceptual Behavior*. Oxford: Oxford University Press.

9

Some Remarks on 'Naturalism' as We Now Have It[1]

Mark Wilson

Many contemporary forms of would-be 'naturalism' have been led up the garden path by Quinean views of 'postulation in physics' that are quite false to the working epistemology of the subject. On this score, neo-Kantianians like Michael Friedman find themselves on prima facie firmer ground. In this note I will offer a brisk survey of the difficulties I have in mind. I should indicate at the outset that I consider myself to be a 'naturalist' (insofar as I understand the term) but believe it must be pursued in a different vein than commonly prevails.

In Paul Benacerraf's well-known 'Mathematical Truth,' it is presumed, without much supporting argument, that the epistemology of how we learn about physical objects is fairly transparent (facilitated, allegedly, by some murky form of 'causal interaction'), whereas the source of our acquaintance with mathematical structures remains a mystery. But, as a matter of descriptive accuracy, this is simply not true: quite commonly we embrace a goodly portion of mathematical doctrine before we infer the existence of quite fundamental physical entities and their properties. (I'll provide a canonical illustration of this dependence below.) Why, then, did Benacerraf presume

[1] After the Birmingham meetings, I thought it might prove useful if chapter 7 were supplemented with a few remarks on the manner in which a profitable 'naturalism' with respect to scientific entities should be pursued. In these considerations, I am grateful to subsequent discussions with Michael Friedman, Bob Batterman, Pen Maddy and many of the participants at the Birmingham gathering.

otherwise? Well, the issue is undoubtedly very complicated, but two sources of the mistake are: (i) a confusion of two senses of 'structure'; (ii) faulty presumptions as to how vital physical quantities find their 'natural names' within a language.

Such assumptions are clearly at work within the recent doctrines that Bob Batterman has characterized as 'mapping accounts of the utility of applied mathematics.' Here it is presumed that the salient 'structures' required in physics can generally be *articulated independently* of a prior acceptance of goodly amount of mathematical doctrine. It is then postulated that mathematics proves itself useful in applied contexts through the existence of *mappings* between such physical structures and correspondent mathematical models. We obtain a 'commuting diagram' picture of applied mathematics as illustrated (see Fig. 9.1), where mathematics entwines itself with physical facts only as an adjacent assistant for the sake of swifter reasoning. (A canonical expression of this view can be found in Hartry Field's *Science Without Numbers*.) From this vantage point, 'naturalism' becomes the thesis that the limited 'physical facts' comprise all that there really is; the chief task is to explain away the extra-physical mathematical auxiliaries.

Sometimes we confront circumstances within applied mathematics that fit this 'mapping' picture fairly well, but we must not suppose that it represents the general situation. To suppose otherwise quickly leads to quixotic projects for disentangling expressive entanglements with mathematics through weird forms of coding or attempts to shelter assertions of substantive 'mathematical content' behind 'I don't believe in it but want to use it' operators. Or through declarations that we are merely talking about 'fictional worlds' when we discuss mathematics. And so on.

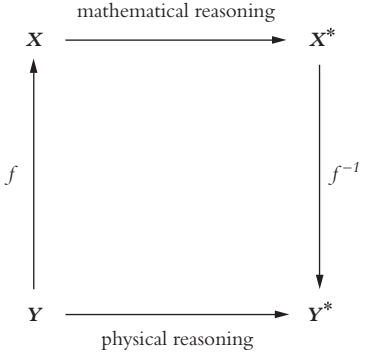

Figure 9.1

Insofar as I can determine, such projects rest upon the tacit presumption that the intuitive notion of 'physical structure' can be neatly captured in the framework of the logician's usual notion of 'structure' <D, ϕ_1, ϕ_2, . . . , ϕ_n>. The latter notion is *linguistically framed* in the sense that we anticipate that a good language for talking about such structures will contain 'kind term predicates' F_1', F_2', . . . F_n' matching the listed ϕ_1, ϕ_2, . . . , ϕ_n traits. It is generally presumed that the other important physical properties of the system can be *generated by first-order logic* from these 'primitive kind terms' F_i. The 'naturalist project' then becomes one of *purging all overtly mathematical relationships* between these in favor of purely logical relations between the physical quantities.[2]

This assumption is quite naïve. In point of fact, the full 'physical structure' of any real life physical system must embrace a huge range of traits (many of them quite important in the natural environment) that lie far outside the orbit of traits obtainable through such limited grammatical constructions (see Fig. 9.2). We require a much different notion of '*reachable property*' if we hope to understand real life physical methodology in an accurate manner. In my opinion, any adequate parsing of the phrase 'physical structure' must embrace this richer gallery of quantities; the

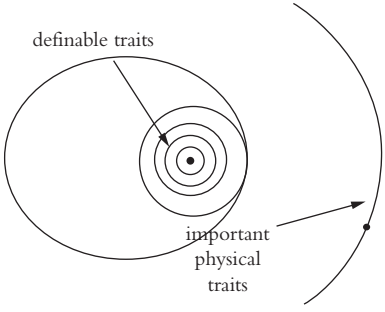

Figure 9.2

[2] Almost always, such reformatory effort is directed only to what the author considers to be 'the fundamental laws or equations of the subject,' in contrast to the mathematical constructions needed to extract useful information from such equations (Green's functions, the Sturmian modes we shall consider in this chapter, etc.) Years ago I asked Hartry Field about such issues and he replied that he 'wasn't concerned with applied mathematics'! I found this response extraordinary in light of the fact that one of the central crises that first brought the explicit consideration of sets to the attention of mathematicians centered upon the role of 'Dirichlet's principle' within the context of the very same equations (Laplace's and Poisson's) that stand at the center of Field's own focus.

notion should not be equated with any <D, ϕ_1, ϕ_2, . . . , ϕ_n> truncation. But as soon as we look carefully into these wider realms of quantities, we find set theory's telltale tools actively at work.

For example, consider the kinds of quantities that physicists locate through *coercive mappings* In essence, a 'coercive mapping' is simply a sequence of questions[3] that can eventually assign the physical system determinate property values at the limiting end of a refining funnel. Indeed, good strategies for playing games like *Clue* or 'Twenty Questions' operate through coercive entrapment of this character: we progressively pose inquiries that ultimately focus upon a final answer ('Colonel Mustard in the dining room with the lead pipe') (see Fig. 9.3). In physics, we generally require an infinity of questions to force out our 'fixed point' quantities.

A prototype for such constructions can be found in Charles-François Sturm's investigations of the 1830s, which construct important descriptive quantities for a wide range of vibratory systems based upon the basic equations they obey. For example, luthiers would like to discover *hidden properties* within their guitars that could help them design better instruments with superior tonal qualities. To this end, they often collect data on what happens if, for example, an upper plate is coated with sand and shaken at various frequencies with various points held fixed. The sand then shifts to places where the wood stays relatively stationary. Why are they doing this? Well, guitars themselves are rather tricky due to their irregular shapes but Sturm demonstrated that allied patterns can be nicely explained by a beautiful set of 'hidden properties' that hide within the vibratory behaviors of rectangular and circular plates (and a wide variety of other configurations possessing a *symmetrical geometry* but of *variable thickness*). He was able to show that such systems are forced by their governing equations and geometry to secretly incorporate *holistic modes of storing energy* that operate independently of one another and which control the behaviors of the plate through a comparatively small range of descriptive factors. If we represent a plate's shifting behaviors as the motion of a point moving within an abstract phase space (left-hand side of diagram), Sturm's hidden quantities allow us to map these complicated movements onto new qualities in terms of which the phase space movements appear more tractable (vide right-hand side of diagram) (see Fig. 9.4). In particular, if

[3] To be more concrete, the answers to such 'questions' commonly assume the form of so-called 'a priori inequalities.'

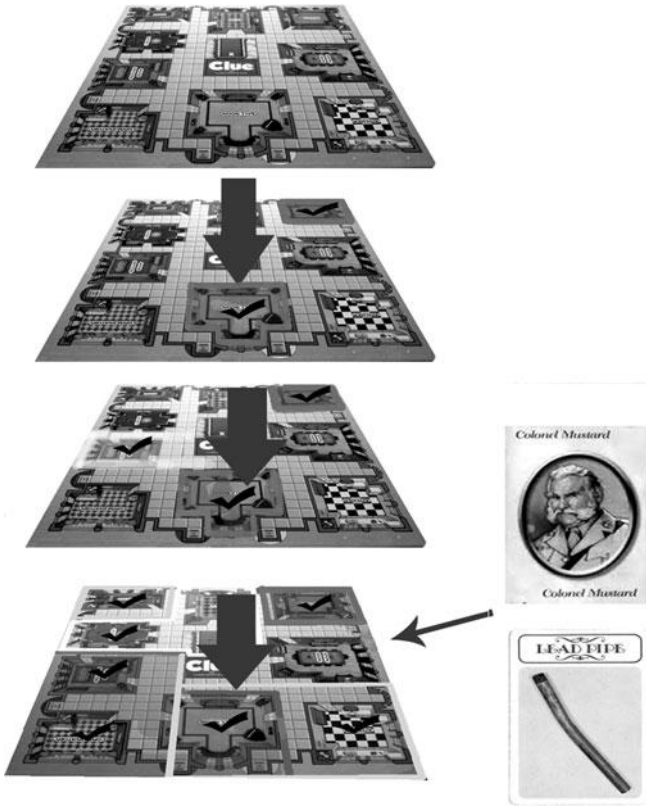

Figure 9.3

only two of these Sturmian modes prove descriptively important, then we have uncovered a map from the plate's complicated-looking spatial behaviors onto a simple movement around a doughnut (more typically, Sturmian mappings carry such behaviors to higher dimensional doughnuts). The utility of these hidden qualities is analogous to the advantages offered in decomposing the movements of a violin string into its overtone spectrum, although Sturm's controlling modes differ from the string's familiar sine-wave quantities due to the shapes of his plates and their variable thicknesses. Once such a decomposition has been uncovered, our abilities to improve the vibratory qualities of symmetrical plates through adjustments to their shapes and thicknesses become greatly enhanced. And this is the design capability that guitar makers hope to emulate.

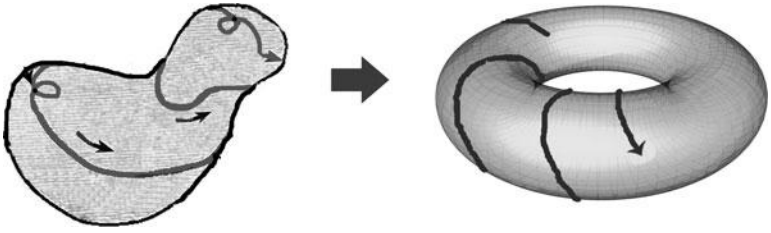

Figure 9.4

The mathematical obstacle that confronts our luthiers stems from the fact that Sturm manages to locate his special decompositional properties via infinitary sequences of coercive mappings that delicately exploit the geometrical symmetries of his special range of systems. Guitars are less regular in their contours and it remains unclear from an applied mathematics viewpoint whether they truly harbor hidden quantities that can offer the same design advantages as Sturm's 'modes.' (It is easier to trap a *squirrel* in a room than a *bat* because the squirrel's movements remain confined to a two-dimensional sheet—similar difficulties obstruct Sturm's 'fixed point' constructions when applied to guitars.) Our instrument makers seek 'controlling properties' analogous to Sturm's but it remains an aspirational hope whether such quantities actually exist.

Observe that Sturm locates his 'hidden' quantities through a strategically informed series of *contractive mappings* from one covering manifold to another, not by locating a preexisting 'kind term' for any of them. Such considerations illustrate what is so deeply mistaken in the 'kind terms' orientation to physical quantities. Recall that such views rely upon the presumption that the important physical properties of a system can be *defined within first-order logic* from 'the primitive kind terms' delineated within the logician's usual notion of 'structure' $<D, \phi_1, \phi_2, \ldots, \phi_n>$. Let's investigate the plausibility of this claim with respect to our symmetrical plaits. Arguably, we can write down the *governing differential equations* for both symmetrical plates and guitars (the equations for both situations are often identical[4]) employing (more or less) only the 'fundamental qualities' anticipated by

[4] It is the manner in which the internal behavior of the plate *interacts with its boundaries* that forces the emergence of the special modes of energy storage that Sturm discovers: the governing differential equation alone cannot guarantee the existence of such quantities.

advocates of the 'kind term' viewpoint. In itself, however, such differential equation information alone is rarely directly informative. Why? In their own rights, differential equations only supply *infinitesimally local information* about their target systems and may not tell us much about how the collection behaves over *finite regions* of space or time. Worse yet, our equations scarcely distinguish a system's varying behaviors over different spans of its career. To appreciate concretely how our target system operates we require charts that cover *finite sections* of its possible behaviors, not simply an infinitesimal snapshot that appears everywhere the same.[5]

With his $<E_i, \theta_i>$ pairs, Sturm happens to have provided *global chart information* with respect to the behavior of his target systems, but more typically applied mathematicians set up smaller descriptive patches based upon the *localized descriptive opportunities* that systems offer (e.g. important physical quantities often capture a system's behaviors quite effectively as long as they remain within a close neighborhood of some equilibrium state). Many recent advances in physics are linked to Sturm-like constructions that assemble their quantities in a more localized fashion.

In either situation, the basic moral remains the same: the mere fact that we can express the underlying differential equations of a system employing only the 'kind terms' anticipated by the philosophers doesn't entail that all of the *vital physical quantities* pertinent to those systems can also be 'defined' from that starting terminology in the usual manner of the logicians. We've seen that Sturm *squeezes in* on his hidden quantities through a sequence of mapping considerations in a clever manner analogous to a good strategy for solving *Clue*, not through producing preexistent 'names' for them. Indeed, in the 1840s Sturm's friend Joseph Liouville proved that such $<E_i, \theta_i>$ traits will rarely be *definable* in terms of our starting $\phi_1, \phi_2, \ldots, \phi_n$ vocabulary in any reasonable sense of 'definability.'[6] Sturm's traits (E_i) connect to our 'fundamental' $\phi_1, \phi_2, \ldots, \phi_n$ quantities as the fixed points of *elaborate nested mappings*, not through simple grammatical relationships (I find it surprising that contemporary philosophers usually ignore these well-known results, especially in light of the historical role that such considerations have played in entangling applied mathematics with set theory,

[5] Analogy: as Alice falls down the rabbit hole, she only knows that she continues to wear the same set of clothes and remains otherwise blind to the passing scenery. Well, the information directly contained within a differential equation is of this same limited character.

[6] Cf. my 'Physical properties' (Wilson 1993).

for 'sets' and 'limits' represent the natural vocabulary for capturing the 'fixed point' constructions to which Sturm appeals).

In other words, the plate's complex behaviors have been projected through an elaborate intervening arena (which we might label 'Sturmland'), from which desirable hidden E_i's eventually emerge as the fixed points of strategically informed mappings. If one inspects the relevant terminologies carefully, one finds that the quantities described therein usually obtain their 'natural names' from the *computational techniques* employed or from the *strategic manner* in which they have emerged from some intervening mathematical arena allied to 'Sturmland' (see Fig. 9.5). Such vocabulary may still sound 'physic-y' to a lay person but their nomenclature does *not* conform to the logistical expectations of the 'kind terms' advocates. Accordingly, the majority of quantities encountered in a physics book do *not* obtain their usual 'names' in 'kind terms' fashion, even if they may appear to an inattentive eye as if they do. There are far more properties in the physical world than dreamt of in the philosophy of 'kind terms' and they are generally interconnected with one another through *structural mappings*, not by grammar.

It strikes me that would-be 'naturalists' need to be more careful in how they conceptualize the phrase 'physical structure': it should not be regarded as some simple counterpart to the logician's $<D, \phi_1, \phi_2, \ldots,$

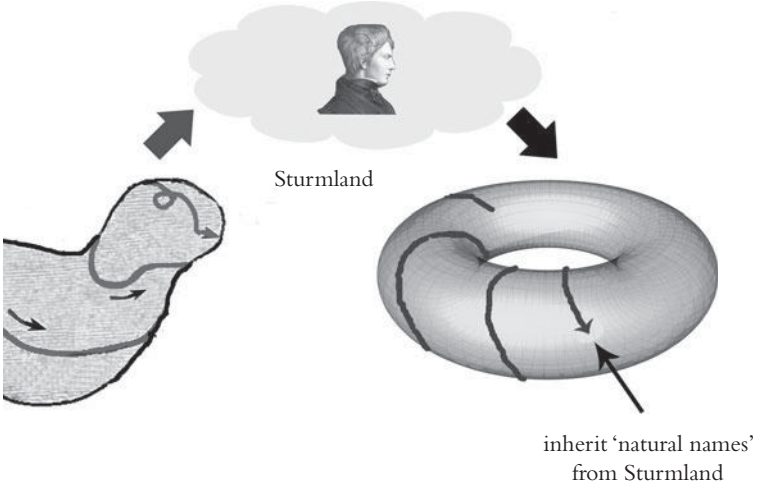

Figure 9.5

$\phi_n>$. Indeed, as Sturm's constructions illustrate, we generally rely upon our *prior sense of mathematical behavior within a set-theoretic vein* in order to flesh out a tolerable sense of the quantities contained in the world based upon a starting knowledge of its governing differential equations and side conditions. Standard Quinean dogma begins with the presumption that 'mathematics is largely posited *for the sake* of the interests of physics.' But such talk of 'posits' severely distorts the epistemological role that abstract mathematics plays within real life science.[7] Without set-theory's assistance, we would remain blind to the world's wider behaviors like Alice in the rabbit hole, for we must exploit the tools offered in orthodox mathematics to flesh out a more robust account of the qualities that the world contains. In these respects, neo-Kantian philosophers working in the vein of Ernst Cassirer can assemble a strong prima facie case that their constructive portrait of scientific methodology accords better with the epistemology of real life practice than do the blinkered 'naturalisms' of the philosophers who follow Quine on 'posits.' Friedman doesn't explicitly incorporate many elements of this complaint within the sketch of modern neo-Kantianism he offers in this volume, but they strike me as potentially valuable supplements to his brief.

In any case, I think that we must have misunderstood our 'naturalism' if we fancy that we are thereby obliged to talk of the world's wide range of quantities (and their entanglements with set-theoretic ideas) only in strange or evasive ways. We require a version of 'naturalism' that is neither eliminativist nor devious about 'sets' in the fashions followed today. I have some vague intimations on how such a 'naturalist' alternative might be developed, but I'll reserve those musings until such occasion as I might develop the relevant considerations more fully and with—I hope!—improved clarity.

References

Benacerraf, B. (1973). Mathematical truth. *Journal of Philosophy*, 70 (19): 661–79.

Batterman, R. (2010). On the explanatory role of mathematics in empirical science. *The British Journal for the Philosophy of Science*, 61: 1–25.

[7] Penelope Maddy develops this criticism of Quine ably in *Second Philosophy* (Maddy, 2007). She eventually advocates a bifurcated 'naturalism' in which set theory has autonomously severed all final allegiance to physical percept. I would prefer a 'naturalism' that demands a somewhat greater loyalty to nature than this.

Field, H. (1980). *Science Without Numbers*. Princeton: Princeton University Press.
Maddy, P. (2007). *Second Philosophy*. Oxford: Oxford University Press.
Wilson, M. (1993). Physical properties. In S. J. Wagner and R. Warner (eds.) *Naturalism: A Critical Appraisal*. South Bend: Notre Dame Press.

10

Causation, Free Will, and Naturalism[1]

Jenann Ismael

> Concepts which have proved useful for ordering things easily assume so great an authority over us, that we forget their terrestrial origin and accept them as unalterable facts. . . . It is therefore not just an idle game to exercise our ability to analyze familiar concepts, and to demonstrate the conditions on which their justification and usefulness depend, and the way in which these developed . . . in this way they are deprived of their excessive authority.
>
> (Einstein quoted in Born, 1971: 159)

The problem of free will arises in a distinctive form in the Newtonian context with the provision of strict deterministic laws of nature that allow our actions to be derived from conditions in place at the beginning of the universe.[2] This is understood as meaning that no matter how much it feels like our actions are flowing freely from our choices, they are in fact causally

[1] I owe a great debt to Michael Mckenna and Keith Lehrer for very helpful comments, and to Dave Schmidtz for the opportunity to present this material in his seminar. Also, a very warm thanks to James Ladyman and Don Ross for including me in the workshop that gave rise to this volume, and the hospitality of the philosophy department at the University of Alabama, Birmingham. And, like all of my work from 2005–10, to the Australian Research Council for support.

[2] This is just one horn of what is known as the Dilemma of Determinism. The problem doesn't disappear if the laws incorporate the sort of indeterminism characteristic of quantum phenomena. The reason is that we want our wills to be the originators of action, not random, spontaneous acts of nature.

determined by facts that were in place before we ever came on the scene. In recent years, defenders of free will in the philosophical literature have focused on the notion of personal autonomy, showing that in paradigmatic cases of willful action there are psychological structures in place that make it right for me to say that action flows from me in a morally relevant sense. The focus turned to psychological requirements for freedom and removing the kinds of impediments we all recognize as undermining the will to act: addiction, delusion, akrasia, misinformation, and coercion. These are the sort of things Denis Overbye is jokingly alluding to when he writes 'I was a free man until they brought the dessert menu around.'[3] These kinds of constraints pose real and substantial threats in everyday life. Freedom from them is not an all or nothing matter and it is not easy. There is a lot of very useful discussion in the moral psychology literature about what free will, understood in these terms, amounts to and how to attain it.

But to someone gripped by the worry that our actions are causally determined by the initial conditions of the universe, talk about morally and psychologically relevant senses of freedom can seem beside the point. We can talk until we are blue in the face about freedom from psychological constraints on action, but if our actions are *causally* necessitated by the earliest state of the universe, we are no more free to act otherwise than a leaf is free to blow against the wind, or a river is free to flow up the mountain rather than down. Of course, we don't typically know the causal determinants of our actions, but that doesn't mean that we are causally capable of acting otherwise than we do.

This is a worry that need not deny that you may act as you choose, but it points out that your choices themselves have causal antecedents. The mere existence of causal antecedents to your choices means that even if you satisfy all of the requirements for personal autonomy, you are causally compelled to act as you do. Here are some expressions of this worry:

> Causal determinism entails that all human choices and actions are ultimately compelled by . . . a complex chain (or web) of causal antecedents. (Ferraiolo 2004: 67)
>
> If determinism is true, . . . all those inner states which cause my body to behave in what ever ways it behaves must arise from circumstances that existed before I was

[3] New York Times [online] <http://www.nytimes.com/2007/01/02/science/02free.html> accessed 17 May 2012.

born; for the chain of causes and effects is infinite, and none could have been the least different, given those that preceded. (Taylor, 1963: 46)[4]

I want to confront this worry head on, starting with the pre-theoretic view of causation, showing how what we've learned from science about the deep logic of causal talk undermines the conception of cause that makes this seem like a worry.

The plan for the chapter is as follows: I will start with the folk notion of cause, lead the reader through recent developments in the scientific understanding of causal concepts, showing how those developments undermine the threat from causal antecedents, and end with a happy complicity of a scientific vision of the world that in no way undermines the happy confidence in one's own powers to bring things about. Then I'll make some methodological comments, using the discussion here as a model for a kind of naturalistic metaphysics that takes its lead from science, letting its concepts be shaped and transformed by developments in science.

The problem of free will is a notorious Gordian knot, a mess of tangled threads that need to be teased apart and addressed in turn. Each one of these threads holds important lessons for our understanding of ourselves and our place in nature. The worry about causal necessitation is only one of these threads, but one that holds some important lessons about the nature of natural necessity.

The evolution of causal notions

The concept of cause has always played a central role in people's thinking about the natural world, but it has undergone a quiet transformation in science that not all philosophers are aware of. Causal thinking arises spontaneously in the child at about the same time that it begins to recognize that it has willful control over its limbs. That process of discovering the robust connections that allow us to act effectively continues into adult life where knowing the causal effects of our actions in the short and long term is essential to effective agency. The most rudimentary ideas about cause and effect have to do with the fact that by exerting various forces and doing work on other material systems, one could make them behave as one

[4] Taylor is not endorsing this view, but formulating what he calls the standard argument against free will.

wants. In this incarnation, causes are closely connected with mechanical concepts like force and work. Science can be seen as developing out of systematization and abstraction of causal thinking, the search for an understanding of the causal relations among events. At first, the notion of cause in science retained its close connection with mechanical ideas. A cause was something that brought about its effect by a transfer of force. But in Newton's physics, this connection is loosened. Physical laws take the form of differential equations that give the rate of change of a quantity over time. Force drops out of the picture. The idea of compulsion or even a *direction* of influence is lost. As David Bohm put it:

> It is a curiously ironical development of history that, at the moment causal laws obtained an exact expression in the form of Newton's equations of motion, the idea of forces as causes of events became unnecessary and almost meaningless. The latter idea lost so much of its significance because both the past and the future of the entire system are determined completely by the equations of motion of all the particles, coupled with their positions and velocities at any one instant of time. (Bohm, 1989: 151)

Russell thought these sorts of relations embodied in dynamical laws were so different from causal relations as traditionally conceived, that it is misleading to think of them in causal terms at all. He gave two main reasons for rejecting a causal interpretation of the dynamical laws. The first reason was what Bohm was remarking on, viz. that causal relations incorporate a temporal asymmetry that dynamical laws do not. They are what are sometimes called regularities of motion that relate the state of the world at one time to its state at others, but there is no direction of determination. There is no suggestion that one state *produces* another, brings it about, or compels its occurrence. It is true that fixing earlier states fixes later states. But it is also true that fixing later states fixes earlier ones. The dynamical relationship is entirely symmetric and incorporates no direction of determination.

The second reason that Russell gave for rejecting a causal interpretation of the dynamical laws was that causal relations hold between relatively localized events at a temporal distance from one another, like the striking of a match and the appearance of a flame, or the turning of a car key and the starting of an engine. The dynamical laws, by contrast, relate only states of the world as a whole (in relativistic physics, spatial hypersurfaces). This is connected to a third difference between dynamical laws and causal generalizations that others have noted. Causal generalizations are imprecise and

defeasible. We all know that to get matches to light and cars to start, all kinds of subsidiary conditions—not usually mentioned and not always explicitly known—have to be in place. Things go wrong, and we don't always know what. Dynamical laws, by contrast, don't allow exceptions. They hold with perfect precision at all times and places. Unfortunately, because they relate global time slices of the world, they are of little use for the purposes of guiding action or prediction at the local scale. Although they are less precise and always defeasible, causal generalizations are much more useful for beings with limited information and local input to the world.

Russell's view in the 1918 paper was that causation is a folk notion that has no place in mature science. As he (famously) says:

> The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm. (Russell 1918: 180)

There are few people who agree with Russell's conclusion, but his remark set off a line of questioning about the status of causal relations. What we've learned, particularly in the last twenty years, about the logic and content of causal claims casts the problem of causal determinants in a rather different light. The modern discussion of causation in the philosophy of science really began with Cartwright's deeply influential and telling criticism of Russell's paper (Cartwright 1979). Philosophers often cite Gettier as a rare example of someone who genuinely refuted a philosophical view, but Cartwright's argument against Russell came quite close. She pointed out that dynamical laws cannot play the role of causal relations in science because specifically causal information is needed to distinguish effective from ineffective strategies for bringing about ends. So, for example, it might be true as a matter of physical law (because smoking causes bad breath), that there is a strong positive correlation between bad breath and cancer. But it is not true that bad breath *causes* cancer and hence it is not true that treating bad breath is an effective strategy for preventing cancer. And that difference—the difference between being correlated with cancer and being a way of *bringing it about*—is not one that can be drawn by looking only at dynamical laws. If one wants to avoid getting cancer, one has to know not simply what cancer is correlated with, but what *causes* it, that is, what brings it about.

There was a lot of handwringing, wondering what causal information adds to dynamical laws. Philosophers entertained probabilistic and counterfactual analyses, and there were a lot of unresolved questions about the

metaphysics of causal relations. In the last fifteen years, there has been a quiet revolution in how we model, understand, and learn about the causal structure of the world. The revolution started in philosophy and computer science, but spread to psychology and statistics where theories of causal learning and statistical inference have made huge steps. For the first time, we have available a comprehensive formal language in which to represent causal structure and which can be used to define normative solutions to causal inference and judgment problems (Pearl 2000).

New insights are emerging into the role of causality in the basic human cognitive functions: decision-making, reasoning, judgment, categorization, inductive inference, language, and learning. Facts about how people informally think, talk, learn, and explain things in causal terms, are formalized and systematized in scientific contexts. There are complementary formalisms (graphs and counterfactual analyses, Bayes nets), but the interventionist account has emerged as a clear forerunner. The formal apparatus gives us ways of rendering the deep causal structure of situations, refines intuitions and gives us positive criteria for making assessments in hard cases.

The central idea of the interventionist account is that causal structure encodes information about the results of hypothetical interventions. 'Intervention' is a term of art that refers to interaction that effectively randomizes the value of a variable, so that it can be treated probabilistically as a free variable. Although interventions are supposed to be defined formally and independently of human action, human actions turn out to be an important class of interventions,[5] and so the role that causal information plays in guiding strategic action (noticed by Cartwright in her objections to Russell) fall very nicely out of the interventionist account of the content of causal claims. We need causal information to decide how to act, because we need to know how our actions will affect the future, quite independently of how they might themselves be affected by their own pasts.

The interventionist account came out of independent work by Glymour's group at Carnegie Mellon and Judea Pearl at UCLA. Pearl's work culminated in his beautiful book *Causality* (2000) and became known to more philosophers through James Woodward's *Making Things Happen* (2003). It

[5] In many interactions, volition-governed action effectively randomizes the effects of external variables, breaking correlations between a variable and earlier parameters.

provides a rich formal framework for representing causal relations in science and makes it easy to express in logical terms what causal information adds to the information embodied in dynamical laws.[6] Dynamical laws tell us how the state of the universe at one time is related to its state at another and they entail a complex web of interdependencies among the values of physical parameters. What causal information adds is information about what would happen if a given parameter were separated out of this web, severing connections with antecedent variables, and allowed to vary freely. The term 'intervention' is used to describe the process of separating of a parameter from antecedent variables and allowing it to vary freely, hence the name and the slogan of the interventionist account: causal information tells us what *would* happen under hypothetical interventions on a chosen parameter A.[7]

The interventionist account captures quite nicely the intuitive idea that causal information separates the *effects* of a parameter from information it carries in virtue of its own connections to causes in the past. Knowing the causal effects of A is knowing how the values of downstream variables are affected by free variation in A. Even though there may be a strong law-like correlation between having breath that smells like cigarettes and developing cancer, bad breath is not a *cause* of cancer if the cancer rate would not be altered by giving people mints to improve their breath. Smoking, by contrast, is a cause of cancer if abolishing smoking would lower cancer rates. Intuitively, causal structure separates what a parameter *does*—namely, the effects that it brings about—from the information it happens to carry about the future in virtue of common causes in the past. Since that is a difference that only shows up when its own links to other variables are severed, and since no variable is ever actually severed from its causes, this extra content can be captured only in counterfactual—or, as interventionists like to say, hypothetical—terms. Think of a newscast in which Obama,

[6] By 'dynamical laws' I will always mean fundamental laws of temporal evolution defined, in the first instance, for the universe as a whole.

[7] It turns out to be subtle to characterize interventions explicitly and there are disagreements among interventionists about the right formal definition. But there is agreement that intervention is an indisputably causal notion. Interventions can be characterized as a special class of causal processes, but not in non-causal terms. So the interventionist account aims for elucidation of the truth-conditional content of causal claims (the inferential relations among them and the conditions and consequences of accepting them), but doesn't reduce causal facts to non-causal ones.

announcing new measures in Afghanistan, follows a weatherman forecasting rain. Firing the weatherman won't ward off rain, but firing Obama would ward off the announced measures in Afghanistan, but the difference—the difference between Obama's connection to the events he announces, and the weatherman's connections to the weather—is a specifically causal one. It cannot be made out in terms of correlations between reports and the events reported.

One of the most interesting consequences of this account from a philosophical point of view is that causal information turns out to be *modally richer* than the information contained in fundamental dynamical laws by which I mean that it has modal implications that strictly outrun the modal implications of the laws. The laws tell us how the state of the world varies over time and hence how history as a whole would vary with different choices of initial state, but they don't have any direct implications about what would happen under free variation of local parameters that occur later in history. And this is the information we need in order to make causal judgments. For, once the initial conditions are given, the counterfactuals we need to assess to capture local causal structure—so-called 'intervention counterfactuals'—are counter*legals*.

It's important that it is understood that there is no logical *incompatibility* between the laws and intervention counterfactuals. The fundamental laws simply *leave out* the information contained in the intervention counterfactuals. They are *silent* about the hypothetical scenarios described in the antecedents of intervention counterfactuals. The intervention counterfactuals contain modal information that *goes beyond* the information contained in global laws; they describe hypothetical cases that the laws do not cover. This is a consequence of the formal definition. It means that for (almost) any set of global dynamical equations involving two or more variables, there will be multiple conflicting accounts of the causal structure of the system that satisfy the equations. These models will preserve the relations among the values of parameters entailed by the laws, but disagree over the results of hypothetical interventions.

A concrete example

To get a more intuitive understanding of how we should think of causal structure, let's get a concrete example in front of us. Consider a fairly familiar kind of mechanical system, say, the engine of a car. If the engine

were its own little universe—that is, if it were closed with respect to external influence—we might be able to come up with global laws that would allow us to calculate its state at one time from its state at any other. If we were just interested in predicting its behavior, this would be all we would ever need. There would be nothing unaccounted for in the model that could make a difference to its evolution, no unexpected contingencies that could divert it from its course, no action on it from the outside and no question of acting on it ourselves because the engine is *not* its own little universe that there are not likely to be such laws. It is embedded in a larger universe and subject to influence from outside. There are breakdowns, unexpected accidents, contingencies that can never be entirely accounted for. Any global rules that describe its behavior are defeasible regularities that come with ill-defined *ceteris paribus* clauses; claims about how it normally behaves, if all goes well.

And it is because we are not simply interested in prediction that even if there *were* such laws, we would need something more. The fact that we interact with the engine means that we are not just interested in knowing how it *does* evolve; we are interested in knowing how acting on it in various ways *would* affect its evolution. This information precedes our knowledge of how it *will* behave because it *guides* our interventions into its behavior. For these purposes, global laws that tell us how the system's state changes if left to its own devices do us little good. We need a working knowledge of the engine. Working knowledge, as we will see, is causal knowledge, and we can get a good understanding of the epistemic and practical significance of causal information by looking at what it takes to get a working knowledge of a system. Suppose Joe is an aspiring mechanic and he wants to know how engines work. It wouldn't be enough to give him a rule that allows him to calculate the state at one time from its state at another. We would tell him what the important components of the engine are and give him rules that told how they work separately and in conjunction to produce the overall pattern of evolution. Very likely, we would start with a diagram of the moving parts, which identified the valves, camshaft, piston, crankshaft, and so on (see Figure 10.1).

And then we would give him separate diagrams for each of the parts that say how their output varies with input. We might start with a model of camshaft that tells him that the camshaft transforms circular motion into a

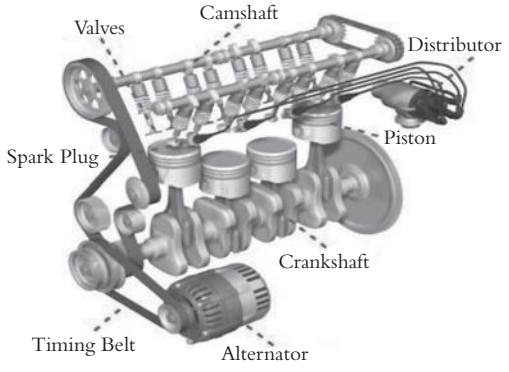

Figure 10.1 The parts of the engine in configuration.[8]

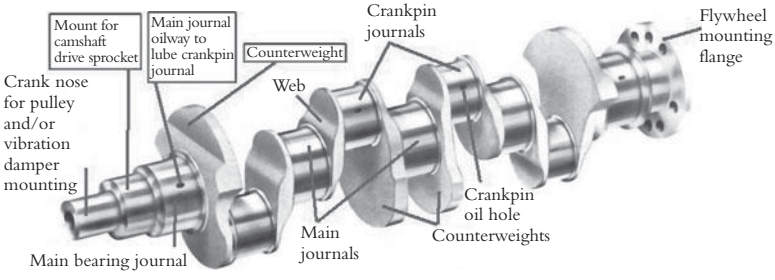

Figure 10.2 The camshaft.[9]

downward force applied to the valves, and how the speed of the motion affects the strength of the force.

And then we might give him a model that shows how the valves work (see Figure 10.2), telling him that each valve utilizes a spring, which will return it to its original position (closed) after the force is removed (see Figure 10.3).

He would also need a model of the piston telling him that the piston moves up and down, opening and closing the intake valve so that at the beginning of the intake stroke, the valve opens drawing the fuel-air mixture

[8] Chongqing CAIEC Machinery & Electronics [online] <http://www.caiecq.com/suzuki-crankshaft-for-sale-p-357.html?cPath=114_118> accessed 17 May 2012.

[9] 2 CarPros [online] <http://www.2carpros.com/articles/how-camshaft-variable-valve-timing-works> accessed 17 May 2012.

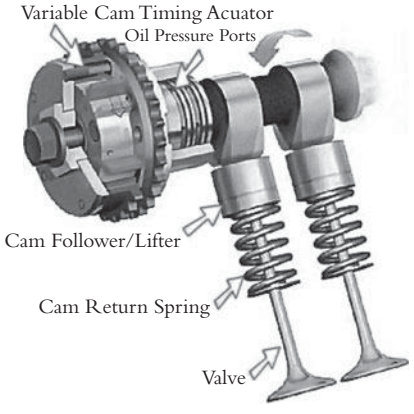

Figure 10.3 The valves.[10]

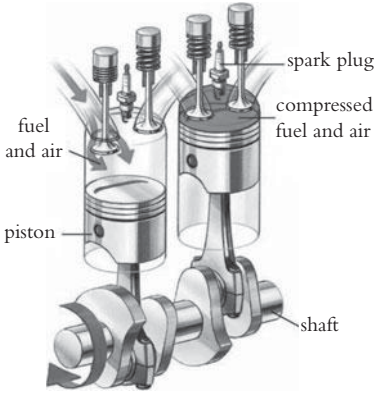

Figure 10.4 The piston.[11]

into the cylinder and when the piston reaches the bottom of the intake stroke, the valve closes, trapping the mixture in the cylinder (see Figure 10.4).

Once he is acquainted with how the parts behave by themselves when their input is allowed to vary freely, they are reassembled in a model that shows how the fixed connections among them constrain their relative motions, that is, how the output of one constrains the input to another. Once he has all of this, he will have what we think of as a working knowledge of the engine in

[10] 2 CarPros [online] <http://www.2carpros.com/articles/how-camshaft-variable-valve-timing-works> accessed 22 May 2012.

[11] Your Dictionary [online] <http://images.yourdictionary.com/piston> accessed 17 May 2012.

the sense that he will know not only how it evolves as a whole, but also how it decomposes into parts, the rules governing the behavior of the parts, and the mutual constraints imposed by their arrangement.

There is a reason for this procedure. There is a reason, that is to say, that we need the kind of working knowledge embodied in the sub-models of the various parts of the system in addition to global laws of evolution.[12] The global laws contain information about how the engine evolves as a unit. The sub-models tell us what would happen if the parts were separated out and their input allowed to vary without constraint. Sub-models tell us what would happen under free variation of parameters whose values are constrained in configuration. The interventionist slogan put this point by saying that the information that is captured in these sub-models tells us what would happen if internal parameters were separated from their pasts and allowed to vary freely. I prefer to put it a little differently by saying we can't in general extract the theory of local subsystems of a larger system from dynamical laws that pertain to the whole, though the point is essentially the same. There are many ways of putting together subsystems to achieve the same dynamics for the overall state. Embedded in a larger machine, the input to each of the parts is constrained by the parts of the engine, so that each now moves within a restricted range and information is lost. That is why in general causal information about local subsystems goes missing when we embed them in a more encompassing model. Variables that were allowed to vary freely in the original model are constrained by the values of variables in the embedding model and so we lose information about what *would* happen *if* they were allowed to vary freely. And again, the more encompassing model isn't *incompatible* with the less encompassing one. It is just silent about a class of hypotheticals that the second one tells us about. If we just look at the way that a system evolves as a whole, we lose information about the compositional substructure that is important to guiding our interaction with it.

The practical importance of causal information is that different ways of piecing together subsystems will have different implications for the results of local interventions. We care about causal information because our interaction with the world takes the form of changes in the values of local variables and acting strategically demands knowledge of how changing the values of local variables affects things downstream. Working knowledge

[12] 'Global' here means 'applying to the engine as a whole.'

matters when we are not just predicting how a system will evolve, but concerned with the question of how setting the values of internal parameters will affect its evolution. For purposes of predicting whether someone who smokes will get cancer, it doesn't matter whether smoking causes cancer or is merely a good predictor, a sign of something else that causes it. For purposes of predicting whether an engine will break down, it doesn't matter whether dirty oil causes, or is merely a sign of, impending engine breakdown. The causal information matters if we are trying to decide whether to quit smoking or whether to clean the engine pipes. It's no use doing either if they are not causally related (respectively) to cancer and breakdown.

This working knowledge of the parts of the world is essential for embedded agents interacting with open subsystems of the world. Philosophers tend to be uninterested in these partial views of bits of the world. They make a lunge for totality,[13] to get a fully encompassing view of the world as a whole as it appears *sub specie aeternitatus*. Most day-to-day science, however, is not concerned with the world as a unit, but is focused on local subsystems of the world, investigating the causal substructure of systems at different levels of description. Science raises and answers questions about what would happen under hypothetical variation of any chosen variable, holding fixed any specified class of internal connections. It was a remarkable discovery when Newton found dynamical laws expressible as differential equations that make the state of the universe as a whole at one time in principle predictable from complete knowledge of its state at any other time. But those laws are of little use to us, by themselves for prediction or for intervention. We don't encounter the universe as a unit. Or rather, we don't encounter it at all. We only encounter open subsystems of the universe and we act on them in ways that require working knowledge of their parts.

The structure of causal sub-models

Any real subsystem of the world can be represented in countless ways: on its own, as part of an indefinite number of larger systems, at different levels of description, holding fixed different elements of internal structure and assuming different external scaffolding. There are different ways of carving things

[13] Milton Munitz uses this phrase to describe what he saw as Einstein's move in arriving at the theory of relativity.

up, and different places to draw a line between fixed and variable structure. We can model the car engine *as an engine*, which is to say, as a collection of macroscopic components bound together into a configuration that restricts relative movement. But we could equally model it as a mass of metal and rubber, holding fixed nothing about how it is arranged, or a collection of microscopic particles not even bound into metallic and rubbery configurations. And we can treat it as we did above, by itself or as part of a larger system. So how does a scientist proceed when he sets out to model a system? He creates a kind of virtual separation of a system from its environment; he puts a kind of frame around it. As Pearl says:

> The scientist carves a piece from the universe and proclaims that piece*: in* namely, the *focus* of investigation. The rest of the universe is then considered *out* or *background*, and is summarized by what we call *boundary conditions*. (2000: 15)

The variables that are included in the model are referred to as endogenous variables. The variables whose values are not included in the model are referred to as exogenous variables. He then specifies a range of values that the exogenous variables can take. This range of variability introduces a degree of freedom into the model.

He then specifies a class of invariants, features of the system that he will hold fixed for purposes of assessing the effects of variation in the values of exogenous variables. These choices affect his conclusions about the system. When we ask how the output of the piston, or the engine as a whole varies with differences in input, it matters very much which elements of internal structure and external scaffolding we hold fixed. By scaffolding, I mean features of the external environment that are not explicitly mentioned in the model but that are crucial to supporting the fixed connections inside the system. So, for example, if we are developing a model of an engine or a building, we make some tacit assumptions about gravity, temperature, and speed. We assume the temperature is not close to absolute zero, that we are not travelling close to the speed of light, and so on. These are accounted for tacitly in the specification of the invariants. If we were travelling close to the speed of light, some of the internal regularities we want to hold fixed would break down. We don't have a well-defined model of a system until we've made choices about the exogenous and endogenous variables and the invariants. Sometimes these choices are tacit, but what we say about the system will depend on them. In modeling the engine above, for example, the conclusions we draw will

depend on whether we hold fixed the internal integrity of its parts and their relative positions, or whether we want the model to cover situations in which those connections are broken.

There are different ways of specifying invariants. We can specify them directly and explicitly by saying 'hold this fixed'. In constructing a model of the engine above, if we wanted to know how the input varies with the output, we might simply stipulate holding fixed the internal configuration of parts and ambient conditions within normal range close to the surface of the earth. But we can also specify them indirectly by specifying the range of applicability of the model, and sometimes that indirect specification is also inexplicit. The normal order of discovery runs in this direction; we choose our endogenous and exogenous variables, make some assumptions—tacit or otherwise—about the range of applicability, and then test for the invariants. We start studying a system across a range of contexts and in retrospect we find that the model doesn't apply in contexts we may only know how to characterize in retrospect.[14] But if we have a fully formulated theory, the theory will tell us what varies with changes in the values of exogenous variables across a specified range of contexts.

Once we've specified the exogenous variables and invariants (or, equivalently, exogenous variables and a range of application), we have a *causal sub-model*. These two things together will induce a separation of the properties of the system into fixed and variable structure. The fixed structure is the internal connections that remain in place across the range of applicability. Every real system will support numerous causal sub-models, each with its own range of applicability and built-in division between fixed and variable structure. And each of these causal sub-models will, in its own way, reveal something about the causal structure of the system.

The important point for our purposes is that sub-models that draw the line between fixed and variable structure in different places are not incompatible, but complementary. What is treated as exogenous in one model is treated as endogenous in another. Aspects of the system that are held fixed in one model may be allowed to vary in another. Each of these sub-models will reveal different aspects of causal structure. Causal sub-models focus attention on interesting quasi-isolated dynamical units that retain enough

[14] Before the discovery of X-rays, we wouldn't have known how to characterize contexts of high radiation. Before electromagnetic theory, we wouldn't have known how to characterize regions of electric and magnetic fields.

internal integrity under action from the outside (at least relative to a range of contexts) to support robust, counterfactual supporting generalizations that can serve as avenues for strategic intervention. The pathways highlighted in these models are defeasible by interaction with other causes, usually supported by unrepresented features of context and often recoverable from lower-level dynamics only by simulation.[15] They play an indispensible role mediating local interaction with the world.

We have an everyday practical need for models that assess the effects of our actions on particular localized subsystems of the world—engines, toasters, computers, and cars—and, more generally, the temporally downstream effects on history of variation in parameters that correspond to decisions, in conditions that leave fixed all of the local structure that isn't directly affected by decision. Those are the models we use in deciding how to act, because we want to know how decisions will affect the history of the world. In constructing those models, we treat our own actions as free variables. And we don't typically just hold fixed the physical laws, we hold fixed all kinds of facts about our environments, all of the known and unknown infrastructure that supports reliable connections among localizable events in our surroundings. But we also have uses for models that assess the effects of differences in family, culture, gender, or early experiences on decisions, models that assess the effects of variation in temperature on signaling in eukaryotic cells, models that assess the effect of variation in inflation rates on employment, ozone on the environment, or sugar on the behavior of naked mole rats.

Relations among causal sub-models.

There are three dimensions along which causal sub-models that represent the same system can differ from one another.[16] First, they can have different *scope*. We saw examples of this already in the relations between the sub-models of engine parts and the wider scope model of the engine in which they were embedded. Second, they can have different *invariants* (or ranges of applicability). In modeling the engine, we tacitly held fixed the internal configuration of moving parts, but we could just as easily have

[15] This is because the robust pathways are emergent structures stabilized in the feedback loop of acting on a system and observing the results of action.

[16] This may not be exhaustive.

included in our range of applicability situations in which the engine came apart, or the parts were laid on the ground apart from one another. The more we hold fixed internally, the less freedom there will be for the system as a whole, that is, the less variation there will be in its overall state. Finally, they can be at different *levels of detail* (we could describe the engine at the macroscopic or microscopic). A system that has only five components at one level of detail may have 10^5 at another. Think of what happens when we look at the engine at the level of microscopic detail.

The causal substructure of a system is only partially captured in any one of the myriad number of different ways of modeling a system, each with its own choice of exogenous variables and fixed set of internal connections. Fully specified, it should furnish the basis for judgments about what would happen under free variation of any chosen parameter, holding fixed any specified class of invariants. When we assess the effects of variation in a parameter, we are usually looking forward in time, but we can also raise questions about the effects of variation in current or future states on the past. There is no intrinsic restriction to localized interventions. We are often interested in assessing localized interventions, but we can certainly raise questions about the effects of variation in complex and distributed events like climate or water mass on global variables. And when we assess the effects of one variable on another, we are almost always tacitly holding fixed aspects of the internal structure of the system, and features of the environment not explicitly mentioned and in some cases not known.

We make all sorts of choices when we construct a sub-model. These choices are governed by the purposes to which the model will be put, and the modal conclusions that we draw about the system are *conditioned* on these choices. What would happen if the value of gravity were allowed to vary within a given range? What would happen if the size of the earth were increased or decreased along a given scale? Or if the wind reached speeds of 200 mph in Los Angeles? Answers to each of these questions depends on which elements of external scaffolding and internal structure we hold fixed and the range of circumstances we are talking about. Those have to be specified before we have a well-defined question.

In *actuality* every variable has a value. Modality is introduced when the value of a variable is allowed to vary over a certain range. The variability in every case is *purely hypothetical*. By letting exogenous variables take a range of values, a model introduces a dimension of hypothetical variation into the state of the world and by specifying a class of invariants, it induces a

division between fixed and variable structure. We can talk about hypothetical variation in the values of variables at any point in the world's history, early or late. There is no fact of the matter about what hypothetically happens. There are only facts about what happens under certain hypothetical suppositions.

Which features of our models depend on the choice of exogenous variables and invariants, that is, which facts are *relative* to the choices that define a sub-model? The asymmetry between the exogenous and endogenous variables introduces a *direction* of influence. We tend to hold the past fixed and let the future vary, because questions about how later states vary with differences in early ones have a special importance for purposes of guiding action, but formally, there is no difficulty in asking about the effects of variation in future states on the past. So the asymmetry between cause and effect, the order of determination, is an artifact of the choice of exogenous and endogenous variables. As Pearl says:

> This choice of [endogenous and exogenous variables] creates asymmetry in the way we look at things, and it is this asymmetry that permits us to talk about 'outside intervention', hence, causality and cause-effect directionality. (2000: 350)[17]

And again,

> The lesson is that it is the way we carve up the universe that determines the directionality we associate with cause and effect. Such carving is tacitly assumed in every scientific investigation. (2000: 350)

And it is not just the *direction* of influence. The *existence* and *strength* of influence depends on what we hold fixed and what we allow to vary. Whether A and B are connected at all and the strength of that connection depends on what we hold fixed. Even the most robust local connection like the connection between smoking and cancer, or the ambient temperature and the level of mercury in a thermometer will disappear if we don't hold fixed a good deal of local infrastructure. These connections are contingent on that local infrastructure and don't hold generally. All of the interesting structure is contained in the local sub-models, and emerges relative to constraints and holding fixed certain elements of internal structure.

So the *direction* of influence is introduced by the choice of exogenous and endogenous variables, and the *strength* of the connection depends on

[17] See Price (2007).

the specification of invariants. The upshot is that the idea that changes in one variable *bring about* changes in another is imposed by these choices. It is not an internal relation between the events. The direction in which influence between A and B runs depends on which you treat as exogenous and which you treat as endogenous. This doesn't mean that there aren't objective facts about what varies with what *relative to a choice of sub-model*. It just means that we have to make the choice of exogenous variables and invariants explicit when we raise such questions. There *are* facts about what varies *under hypothetical variation in X when Y is held fixed*. There is no fact about what is *really* or *absolutely* fixed, or what is *really* or *absolutely* variable. The variation in question is in every case hypothetical, and what varies with it is always relative to a choice of invariants or range of applicability.[18] This is all that realism about causal structure requires. In many ways, it is an entirely commonsensical picture that is very natural for the people with hands-on experience investigating causal structure.

The Pearl Inversion: A reorientation in our understanding of modality

I've gone into this in some detail because it is important to see that the idea that the wide-scope models replace, subsume, or reduce the narrow-scope sub-models that they embed is a mistake about the logic of these relations.[19] Wide-scope models don't override, displace, or compete with narrow-scope models. Models that draw the line between fixed and variable structure in different places complement one another, revealing different aspects of the world's modal substructure. The narrow scope models contain causal information that the wide-scope models simply leave out. The rules that govern the behavior of the parts of a system independently are *modally*

[18] Again, it is worth emphasizing that the invariants are usually specified indirectly by fixing a range of applicability and *discovering* what remains fixed across that range. This is why causal structure is a discovery rather than a stipulation. Perhaps the right way to say it is that if we fix the range of applicability, the world will fix the invariants. But we can always change the range of applicability, and when we do, the modalities will change as well.

[19] What makes the discovery of causal structure possible, where it is possible, is the existence of processes that have the effect of randomizing the values of exogenous variables. We can 'carve off a piece of the world,' and actually study the effects of random variation in exogenous variables. The results obtained will be only as good as our capacity to randomize and there are complicated practical issues about how to achieve and ensure randomization, but there is nothing suspicious about the status of this extra modal information.

stronger than the rules that govern them in configuration in the sense that they cover situations that don't arise in configuration. They have counterfactual implications that outrun the counterfactual implications of the rules that apply to the configuration. So far, this is just an observation about the logic of the relations among sub-models, that is, an artifact of the embedding relation, a consequence of the fact that when we embed a narrow-scope model in a wider scope one in a manner that constrains the values of variables that are treated as free in the former, we lose modal information.

But it carries an important lesson. There is an unexamined presumption in the philosophical literature that modal truths must be derivative of global laws. But on looking through the literature, I could find no good argument for this presumption. There is a mereological principle that all of the *categorical* facts in a narrow-scope model are included in a wide-scope model that embeds it. But I see no modal analogue. And reflecting on the lessons from the engine militate against it. If we used to presume that the rules that govern the parts of a complex system must be derivative of rules that govern the whole, we now see that the order of priority runs in the other direction. Rules that govern the whole are derivative of rules that govern the parts. We start with the basic building blocks with a great deal of freedom of movement and build up more complex systems by restricting their relative motion. A full account of the modal substructure of a complex system would have to recover the rules that pertain to the parts as well as the whole.[20] This is Pearl's attitude announced in the preface to causality, as he says:

> [I used to think that] causality simply provides useful ways of abbreviating and organizing intricate patterns of probabilistic relationships. Today, my view is quite different. I now take causal relationships to be the fundamental building blocks both of physical reality and of human understanding of that reality, and I regard probabilistic relationships as but the surface phenomena of the causal machinery that underlies and propels our understanding of the world. (2000: xiii–xiv)

In my mind, this is a very deep insight. The struggle to derive causal information from global dynamical laws prompted the reflection that led to it, but it is a conclusion that has consequences for how we think about modality quite generally. Realism about causal structure is realism about

[20] It would allow the assessment of intervention counterfactuals for any choice of exogenous and endogenous variables, relative to any class of invariants.

the modal substructure that underlies regularities in the actual pattern of events. This modal substructure underwrites, but is much richer than, the modal claims embodied in global laws.

Let's recap, separating what we say about the logic of causal claims from any claim about the metaphysics. We say that causal information is information about rules that describe the behavior of the parts of the system, individually and in configuration, where 'rule' is being used here in a neutral, formal way, to mean 'counterfactual supporting regularity.' Rules that pertain to the system as a whole, 'in full assembly,' are stronger than those that pertain to the parts individually. The causal structure of the system captures all of the information about the rules that describe the behavior of the parts individually and in configuration. Causal realism is realism about the relativized counterfactuals: what would happen to A under free variation of B, holding fixed a specified class of invariants? The local rules (rules pertaining to open subsystems interacting with an environment)[21] are prior. The global rules are derivative, obtained by adding constraints on configuration of parts. Experimentally, we discern causal structure by studying the parts separately and then in configuration, holding fixed different aspects of internal structure and external scaffolding. Regularities of various kinds can provide clues to causal structure, but the possibility of intervention and controlled experiment is essential.[22]

If there is no conflict between sub-models that draw the line between fixed and variable structure in different places, how do we choose which sub-models to employ *in situ*? The choice is problem driven, and context dependent. Practical considerations define the scope of a model. We construct sub-models for purposes at hand, choosing what to include, what to hold fixed and what to vary in a manner dictated by the problem context. In a laboratory setting, we're usually dealing with an object system and treating a collection of identified variables whose values we are interested in and have ways of controlling as exogenous. When we are

[21] There is a little terminological ambiguity here. I follow practice in the interventionist literature of calling systems 'open systems,' meaning systems that are subject to external action. In thermodynamics, the notion is used in a slightly different manner, but it won't matter here.

[22] I haven't said anything about the metaphysics, but we get a *prima facie* intuitive understanding from the engine example of how modal facts can be grounded in the concrete material configuration of parts of which a system is composed, each with its own intrinsic range of motion, but bound into a fixed configuration across a range of circumstances.

trying to solve a decision problem we treat our actions as free variables and assess the expected outcomes of different choices. But we model ourselves, too, as individuals and as a group, treating our own actions as endogenous variables and seeing how they respond to differences in others. Recognizing these kinds of patterns can help us understand the cultural, environmental, and sub-personal forces that shape our own behavior.

In decision contexts, what we hold fixed is a partly causal question that depends on what we can expect to be fixed in the hypothetical circumstances in which the choice will take place. And this brings out why decisions are hard. If I am wondering whether I should move to Miami for part of the fall, I can't just hold everything that is the case now fixed, I have to know how the weather will be at the time and whether my children will still be in school. And I have to judge how my priorities and feelings will have evolved in the meantime. Judging that I will want to eat in an hour, even though I'm not hungry now, is a relatively easy call, but judging whether I'll want to be with a partner twenty years down the road, or how I'll feel about having children if I take the plunge now, are not. There is no simple recipe for making these judgments. They are causal judgments, but ones that demand self-understanding and practical wisdom beyond mere scientific knowledge. And this is to say nothing about all of the hard moral questions. What do we hold fixed for the purposes of assessing responsibility for events? What should we hold fixed and allow to vary for the purposes of assigning praise and blame? What should we blame for a car accident that injures a pedestrian? Do we blame the driver's slow reflexes? Those caused the crash only if we hold fixed the weather, the worn tread on the tire, the other driver's poor eyesight and the slow pace of the pedestrian? Who is to be blamed for the Gulf oil spill or the Japanese nuclear disaster? In systems that contain multiple interacting human and non-human components, there is no absolute answer to the question 'who, or what, is to blame?' Where the locus of control lies will depend on what we hold fixed and what we allow to vary.

The problem of causal antecedents defused?

Let's revisit the problem of causal antecedents with this understanding of the deep logic of causal judgments in hand. To get the purported conflict with free will going we are invited to see action in the context of wider embedding models—psychological models, neuroscientific models,

sociocultural models—in which action appears constrained by antecedent variables. The worry about physical determinism is the most extreme version of this sort of model in which action is not just *constrained* but *determined*, and determined by variables that we can trace to the beginning of time. These strategies for undermining freedom purport to take the reins out of our hands. The wide-scope view of action they provide is supposed to override the more limited view we adopt for decision and show us that treating our own actions as free variables is deluded. But that is a mistake. There is no incompatibility between the wide-scope and narrow-scope views of action. Our actions appear in multiple models, sometimes as exogenous, sometimes as endogenous. Whenever we model a subsystem of the world—whether it is a car engine, or a human being—we can always widen our view to attain a more encompassing perspective, and wide-scope models will have a set of possibilities that is typically constrained relative to the narrow scope model.[23] But there is no more conflict between these models than there is between the view of a building from close-up and the view from a very great distance.

The choice between models is pragmatic rather than descriptive. Where our choices make a difference to what happens, we use models that treat our choices as free variables so that our choices can be guided by their projected outcomes.[24] In that context, the hypothetical possibilities that matter are the ones that correspond to different choices and hold fixed features of the world that are invariant under choice and expected to obtain in the context of action.

Science, metaphysics, and common sense

I have been suggesting that the concept of causation formalized in work on the logic of causal modeling defuses the apparent conflict between the first personal view of action in decision contexts and the wide-scope view of action as it appears *in global models*, making the choice between models a pragmatic one and defusing the idea that causes exercise a freedom-undermining compulsion over their effects. Causal thinking is a tool, a

[23] The only exception is the degenerate case in which the newly exogenous variables are random relative to the old across the relevant range of applicability.

[24] For more on the logic of the choice context, see my 'Decision and the Open Future,' forthcoming.

cognitive strategy that exploits the network of relatively fixed connections among the components of a system for strategic action, *not* a coercive force built into the fabric of nature that imbues some events with the power to bring about others. Far from undermining freedom, the existence of causal pathways is what creates space for the emergence of will because it opens up the space for effective choice.

There is an ongoing discussion of causal notions in analytic metaphysics that proceeds by eliciting intuitions about hypothetical scenarios in an attempt to systematize everyday intuitions about when it is right to say that A caused (or was causally implicated in the occurrence of) B. Typical discussions start with a paradigm example (e.g. Suzy throws a rock that causes a window to break). An analysis is offered (e.g. c is a cause of e exactly if e wouldn't have occurred if c hadn't occurred). Then problem cases and apparent counterexamples are entertained to try to refine the analysis (e.g. pre-emptive causes, preventive causes, and cases involving overdetermination).[25] A charitable construal of the goal of these discussions is to provide an analysis of the everyday notion of cause, or to explicate the folk theory that underlies everyday causal judgments. The methodology and the goal in what I've been doing are both very different. There is no consulting intuitions or armchair reflection on the concept of cause. And what is offered is not an *analysis* of the everyday notion. It doesn't purport to give the meaning of the word 'cause,' or the content of anybody's causal concept. It is, rather, as an explication of the scientific notion of cause, which is a refinement and generalization of everyday causal notions. The everyday concept of cause is a clear ancestor, reconstructed in retrospect as a crude, folk version of the developed notions, less precise and less general because conditioned on contingencies about our selves and the context of use.

Science is in the business of giving an account of the universe and our place in it, and to that extent, it is a form of metaphysical inquiry. How should someone looking to science for answers to metaphysical questions proceed? Let the science develop and see where it leads. Physics seems to speak of matters that are thoroughly familiar. But as it increases the scope and depth of its descriptions, it transforms and generalizes those notions. Think of how our concepts of light, matter, and sound have been transformed by science, to say nothing of space, time, and motion. That is what

[25] See Collins, Hall, and Paul (2004) for a good sample of this literature.

has happened with cause. The everyday notion of an intrinsic, compulsive force between natural events has been quite thoroughly transformed. What science tells us is really there—or if you like, 'there anyway'—at the fundamental level as modal substructure is not intrinsically directed, and not an internal relation between pairs of events. The structure captured in causal models identifies relations among variables always relative to some class of invariants, with a direction imposed by choices of endogenous and exogenous variables.

The refinement and generalization that causal notions undergo in the hands of science is characteristic of the scientific development of everyday notions. It replaces the ideas of common sense with more exact and general notions. Pragmatic, contextual sensitivity is eliminated in favor of explicit relativization. Features of the pre-theoretic concept that have an analytic or *a priori* character are often revealed as conditioned on contingencies from a more fundamental perspective.[26] The transformations tend to be conservative of the important core of our practices with the notions, but they can be ruthless with the metaphysical pictures that often accompany those practices.[27]

I don't doubt that there are conceptions of causal notions in the metaphysics literature that make it out to be the kind of asymmetric compulsive force writ into the fabric of nature that would undermine personal freedom, and even that such conceptions may come closer to capturing the core of the folk conception of cause. But the naturalistic metaphysician takes her concept of cause from science, and it is not fair to impose prescientific ideas of causation in an argument that science makes no room for free will. Not only is there room for personal freedom in a fully articulated naturalistic picture of the world, the naturalistic picture gives us an understanding of ourselves and our place in nature that is much richer and more faithful to the experience of agency than the naïve metaphysical images that accompany common sense.

[26] In this case, the direction of influence turns out to be imposed by the choices of endogenous and exogenous variables, and the specification of invariants that is usually left tacit in the everyday application of those notions (in the form of ill-defined *ceteris paribus* clauses) is made explicit and systematized.

[27] In this case, the practices we want to preserve are the practice of treating choice as a free variable in decision, and that of holding ourselves and others responsible for our choices. I have focused on the first. The second takes separate argument.

Of course, a full naturalistic resolution of the conflict between freedom and causal determinism will require parallel developments in our understanding of freedom. Dennett has been the most outspoken explicit naturalistic revisionist about the common sense notion of free will and *Freedom Evolves* (2003) is a masterful exercise in the style of naturalistic metaphysics that I am recommending. Galen Strawson has complained that Dennett's view doesn't vindicate pre-theoretic conceptions of freedom. He writes:

> [The kind of freedom that Dennett argues we have] . . . doesn't give us what we want and are sure we have: ultimate, buck-stopping responsibility for what we do. It allows, after all, that the whole course of our lives may be fixed down to the last detail before we've even been conceived. (Strawson, 2003: 2)

He may be right. People not only have ideas about causation that give it the kind of coercive unidirectional force to undermine their sense of freedom, but ideas about what freedom amounts to—'ultimate, buck-stopping responsibility'—that conflict with the existence of wide-scope models in which our actions appear causally determined by antecedent variables. But Dennett's reaction to Strawson's complaint—'so much the worse for pre-theoretic conceptions'[28]—is, from the point of view of the style of naturalistic metaphysics that I am recommending, exactly appropriate. The theory of general relativity doesn't preserve pre-theoretic intuitions about space and time, but so much the worse for those intuitions. Someone looking to preserve pre-theoretic ideas will find little satisfaction in science. But someone looking to explore what kind of freedom a naturalistic picture of ourselves allows, will find reinforcement of her practices of choosing and valuing, placed on stronger foundations and with more explicit positive criteria for making informed choices.

Connections

There is little to disagree with from my perspective in the rest of *Freedom Evolves* but the chapter on causation could be strengthened in a way that takes advantage of formalism and developments that Dennett didn't have a chance to exploit. In the 'Notes on Sources and Further Reading,' Dennett remarks that he discovered Pearl's book only while preparing the final draft, saying:

[28] These are my words, not his, but a fair paraphrase.

> Judea Pearl's *Causality* . . . raises questions about [my way of putting things], while opening up tempting alternative accounts. It will be no small labor to digest these and, if need be, reformulate our conclusions . . . this is work for the future. (2003: 95)

It's work that I've tried to do here. The discussion can also, however, serve as a model of a style of naturalistic metaphysics that picks up some of the themes in *Every Thing Must Go*. One of the complaints that critics have made about the book is that they don't provide a clear positive alternative to analytic metaphysics; a clear idea of what naturalistic metaphysics might be like. Cian Dorr, for example, writes in *Notre Dame Philosophical Reviews*:

> The authors' self-proclaimedly 'scientistic' view of the supremacy of the sciences among forms of human inquiry makes it prima facie puzzling how there could be room for a field called 'metaphysics' that is not itself a science and is nevertheless not a waste of time. One traditional picture imagines metaphysics as standing to physics as physics stands to biology—investigating a proprietary realm of facts that cannot even be expressed in the vocabulary of physics, and that provide 'foundations' for physics in the same sense in which physics provides foundations for biology. The authors are skeptical about whether there is any such realm of facts, and thus whether there is any legitimate enterprise of investigating them. What else might 'metaphysics' be, then? (Dorr, 2010)

I certainly agree with Ladyman and Ross that there is no science-independent standpoint from which to do metaphysics. I don't know what naturalistic metaphysics is if it is not simply to bring everything we learn from science—about the world and about our place in it—to bear on philosophical problems, and to grant the scientific view authority over any pre-theoretic intuitions we might have about how things are. I have tried to give concrete content to one way of doing that here, but I don't think there is likely to be any simple template for naturalistic metaphysics or any simple way of demarcating it from science.

Even this weak view of naturalistic metaphysics has opponents, if this closing remark from Dorr's review is anything to go by:

> *Every Thing Must Go* seems likely, if it has any effect at all on analytic metaphysicians, merely to confirm a few more of them in their impression that no one has yet shown how developments in the sciences might be relevant to their concerns. (2010, last line of review)

I'm not sure whom he's speaking of or whether he counts himself in that camp. Most of the analytic metaphysicians I know are engaged with and interested in developments in science. But if his own remark has any effect it will most likely be to confirm philosophers of science in an increasingly pronounced distaste for unscientific metaphysics.

References

Bohm, D. (1989). *Quantum Theory*. New York: Dover.

Born, N. (1971). *The Born-Einstein Letters*. New York: Walter.

Cartwright, N. (1979). Causal laws and effective strategies. *Noûs*, 13: 419–38.

Hall, N. and Paul, L. (eds.) (2004). *Causation and Counterfactuals*. Cambridge MA: MIT Press.

Dennett, D. (2003). *Freedom Evolves*. New York: Viking Press.

Dorr, C. (2010). Review of *Every Thing Must Go: Metaphysics Naturalized*. *Notre Dame Philosophical Reviews* [online]. Available: <http://ndpr.nd.edu/news/24377/?id=19947> accessed 17 May 2012.

Ferraiolo, W. (2004). Against compatibilism: Compulsion, free agency and moral responsibility. *Sorites*, 15: 67–72.

Ismael, J. (forthcoming). Decision and the open future. In A. Bardon (ed.) *The Future of the Philosophy of Time*. Oxford: Oxford University Press.

Ladyman, J. and Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Pearl, J. (2000). *Causality*. Oxford: Oxford University Press.

Price, H. (2007). Causal perspectivalism. In H. Price and R. Corry (eds.) *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Oxford University Press.

Russell, B. (1918). On the notion of cause. In *Mysticism and Logic*. London: Allen & Unwin.

Strawson, G. (2003). Evolution explains it all for you. *New York Times* [online]. Available: <http://www.nytimes.com/2003/03/02/books/evolution-explains-it-all-for-you.html?src=pm> accessed 17 May 2012.

Taylor, R. (1963). *Metaphysics*. Englewood Cliffs, NJ: Prentice-Hall.

Woodward, J. (2003). *Making Things Happen*. Oxford: Oxford University Press.

This page intentionally left blank

Index

a posteriori
  content of science 15, 41–2, 93–4
  methods of science 33–4, 38, 42–6
a priori
  constitutive 10–11, 35, 37, 114–15, 184–6, 193
  framework 10, 18–19, 182–3
  functional 35, 37
  knowledge 2, 5, 16, 60, 63, 114
  principles 10, 34–6, 118
  reasoning 3, 51, 54, 57, 60, 63, 66, 69–70, 80, 94, 112
  relativized 5, 11, 19, 154–5, 158–61, 163, 175, 180, 184–6
  theorizing 32–4, 37–8, 41, 46–7
  truths 3, 10, 60, 95
abduction 142–3
abstract entities 52–3, 102–5, 120, 145, 179, 180 n. 28, 193
aether 35, 37, 131, 162, 194–5
affordances 102
Albert, David 134, 136
antecedent problem for free will (*see* free will)
Anthony, Louise 5
anthropocentrism 71 n. 34, 112 n. 3, 117, 128
anthropology
  naïve naïve 98
  sophisticated naïve 98–106
  auto- 17, 23–4, 98–106
Aristotle 112
artificial intelligence 66, 97–8
Austin, John Langshaw 153 n. 4, 154, 175,

Batterman, Robert 120, 146, 180 n. 28, 199
  The Devil in the Details 120, 146, 180 n. 28
Bayes nets 213
Bealer, George 3–4, 59 n. 17, 66
bees 64
Bell inequalities 136
Benacerraf, Paul 198–9
Bennett, Charles 108 n. 1
Berkeley, George 175
blobject 121
block universe 75
Bobrow, Daniel 98
Bohm, David 55, 132, 134, 135 n. 10, 137, 211
  Quantum Theory 211
Bohr, Niels 112, 125, 130, 134, 135–6
bounded rationality 67 n. 28
bridge laws 63 n. 25
Brown, John Seely 98

cahoots 102
calibration 71
caloric 124–5, 131
Canberra project 17
Carnap, Rudolf 6–7, 10, 14, 72, 101, 118, 119 n. 5, 123, 152, 153 n. 5, 158, 184, 185 n. 4, 187
Cartwright, Nancy 24, 121, 122, 212, 213
Carus, André 153 n. 5, 178 n. 26
Cassirer, Ernst 184–5, 188, 205–6
Cauchy, Augustin-Louis 165, 180 n. 28
causation 20, 208–10
  and counterfactuals 212–13, 215, 228, 231
  and intervention 212–15, 220
  and modality 214–15, 218, 224, 226–8
  as compulsion or force 20, 210, 232–3
  causal antecedents (problem of) (*see* free will)
  causal decomposition 215–20
  causal information 212–20, 226–8
  causal necessity 209–10
  causal sub-models 220–5
  folk conception.of 20, 21, 24, 88, 210–12, 231–3
  graphical approaches to 20–1, 22, 24, 213–15
  scientific conception of 31–2, 120–1, 123, 128, 144, 231–2

causation (*cont.*)
temporal asymmetry of 210, 211, 225, 232–3
center of gravity 104–5, 183, 193
center of narrative gravity 104
Chakravartty, Anjan 15–16, 21–3
Chalmers, David 55, 63–4, 66
chess/chmess 96, 120
cicadas 145–6
coercive mappings 200–5
colour terms 12
Columbus, Christopher 100
computation 71–2, 74, 120, 165–9, 179–80
computational science 64–5
computerised tomography 65
concepts 5–9, 11–14, 18–19, 23–4, 34–6, 37–8, 62–6, 76, 114, 151–64, 174–5, 177–80, 185–6, 188, 190–4, 230–3
classical and anti-classical conceptions of 6–7, 11–12, 18, 23–4, 152–4, 157–9, 164, 171–2, 190
conceptual analysis 1–3, 17, 23–4, 32, 51–2, 59, 62–6, 76, 112, 176
conceptual conservatism 110, 114–15, 118–19, 124, 143–4, 153 n. 4
philosophy as 'critic of concepts' 152, 155, 158–9, 160–2, 163–5, 174–5, 180
confirmation 118, 141
conservatism 7, 110, 112, 113–15, 118–19, 124, 143–4, 153 n. 4 (*see* domestication of science)
constitutive principles 10–11, 35, 37, 114–15, 161, 184–6
constitutive theory (Einstein) 195
continuum mechanics 173, 191
contractive mapping 200–5
conventionalism 35–6, 72, 118
correctness proof 168–9
counterintuitiveness 64, 97–8, 99

d'Espagnat, Bernard
On Physics and Philosophy 112 n. 3, 136
Day, Timothy 7
De Caro, Mario 2–3
de Kleer, Johan 98
decoherence 136
Dennett, Daniel 13–14, 17, 20, 23–4, 68 n.31, 108, 233–4
Brainstorms 105
Freedom Evolves 20, 24, 233–4
lost sock center 104–5
determinism 37–8, 56, 74–5, 125, 131, 137, 144, 146, 208–12, 229–30, 232–3
dilemma of (*see* free will)
Deutsch, David
The Beginning of Infinity 17–18, 111–48
dints 102
dollars 104
domestication of science 112, 143–4
Dorr, Cian 147, 234
Duhem, Pierre 38 n. 4, 187, 189, 191, 192 n.15
Dummett, Michael 170, 177
Dupré, John 122

Ebbs, Gary 188 n. 9, 188 n. 10
Einstein, Albert 10–11, 19, 114, 118, 125, 136, 139, 146, 155, 162–3, 183–6, 193–5, 208, 220 n. 13 (*see* relativity, geometry, Euclidean and non-Euclidean)
electrodynamics 162, 183–5, 194–5
emergence 38, 51 n. 1, 120–1, 133, 135, 223 n. 15
empiricism 39, 48, 51, 70–2, 117–18, 119, 146–7
endogenous variables (*see* variables, exogenous and endogenous)
engine 215–18, 221–2, 223–4
Enlightenment 113, 116, 118, 124, 139, 143, 148
optimism 139, 147, 148
epistemic augmentation 67–8
epistemic extension 64–8
epistemic risk 15–16, 45–9, 58, 69–72
epoché (*see* Husserl, Edmund)
ether (*see* aether)
Euclidean geometry (*see* geometry, Euclidean and non-Euclidean)
Euler's method 167–70, 179
Everett, Hugh 120–1, 122, 124, 125–6, 132–8, 142
exogenous variables (*see* variables, exogenous and endogenous)
experiential distance 44–9 (*see* epistemic risk)
experimental philosophy 99–100

explanation 6, 49, 88–90, 116–17, 119, 120, 121, 125, 130–1, 139–43, 145–6
  explanatory reach 116–7, 125, 127–8, 138, 142 (*see* prediction, unification)

facade (*see* theory facade)
fallibilism 76, 118, 123, 130
falsification 45, 116, 141
fatigues 105–6
Fechner, Gustav 144–5
Feynman, Richard 72 n. 35, 156, 165, 169
fiction 124, 132, 160, 199
Field, Hartry 53 n. 7, 180 n. 28, 199, 200 n. 2
  Science Without Numbers 53 n. 7, 180 n. 28, 199
Fine, Arthur 55 n. 10, 73, 125
Fodor, Jerry 159
folk theories 17, 64, 97–99, 210–11, 212, 231–2
  folk physics 97–8, 99, 212 (*see* causation, folk conception of)
free creativity 154, 155–6, 160, 163, 178–9
free will 99, 208–10, 229–30, 232–3
Frege, Gottlob 58 n. 15
Friedman, Michael 5, 6, 7, 9–12, 18–19, 22, 23, 35, 114–15, 118, 119, 123, 151–5, 158–64, 174–5, 177–9, 180, 198, 206
  Dynamics of Reason 9, 10–11, 18–19, 35, 114–15, 152–80, 185, 186–8
  Kant and the Exact Sciences 9–10, 183 n. 1
  Logic, Mathematical Science and Twentieth-Century Philosophy (LMS) 151–3, 158, 159, 164, 175, 178–9
  Reconsidering Logical Positivism 6, 9, 10, 114, 118, 185 n. 4
fundamental physics 13–14, 51 n. 1, 56–7, 82, 86 n. 10, 89–93, 94, 108–9, 120, 132, 133, 135, 137–8, 146–8

Galchen, Rivka 136
geometry, Euclidean and non-Euclidean 10–11, 118, 123, 160–1, 174, 182, 184, 193 (*see* thought experiments)
Gettier, Edmund 64, 212
Ghirardi-Rimini-Weber theory (GRW) 132, 134, 136
Gibson, J. 102
Glymour, Clark 7, 20, 213
GOFAI (*see* artificial intelligence)
Goldman, Alvin 64
grounding 8–9, 15–16, 39, 40–9, 136
guitar 201–3

Hacking, Ian 36, 130, 142–5
  The Taming of Chance 142–5
Hall, Ned 21
Harman, Gilbert 143
Hawking, Stephen
  The Grand Design 72 n. 35, 110
Hayes, Patrick 97–8, 99
hazy holism 7, 159, 188
Heaviside, Oliver 171
Hebb, Donald 97
Heisenberg, Werner 124
Helmholtz, Hermann 11, 152, 153 n. 4, 153 n. 5, 155, 174, 186, 195
Hempel, Carl 27, 40, 141
Hertz, Heinrich 153 n. 4, 156–7, 161 n. 14, 163, 166 n. 18, 172, 176, 187, 191
  Principles of Mechanics 156, 157, 176
Herzberg, Frederik 73
heuristics 37–9, 43
higher-order truths 95
Hilbert, David 73, 157 n. 8, 183–4, 186, 188–9, 191–2, 193
  Hilbertian axiomatics 157 n. 8, 183–4, 188–9, 191–2, 193–4
holes 102–3
Horgan, Terence 121
human scale 68–9, 116–17
Humean supervenience 4, 56–7
Humphreys, Paul 4, 5, 7–8, 11, 16, 21–2, 23
Husserl, Edmund 99
hypothesis (*see* abduction, Peirce, Charles Sanders)

idealisation 60, 62, 66–8, 76
implicit definition 10–11, 62–4, 114, 153–4, 156, 157–8, 160, 173, 183–6, 188–9, 191–3
indeterminacy of translation 6 n. 2, 105
incommensurability 185, 186, 188
induction 68–70, 117–18, 123, 141–2
  inductive risk 69–70 (*see* epistemic risk)

induction (*cont.*)
inductivism 141–2
inertial frame 114, 163–4, 183–6, 193, 195
inference to the best explanation 7, 49, 120, 143
information and information theory 14, 17, 29, 51 n. 1, 106, 108, 120, 121
instrumentalism 22, 39, 119, 124–8, 129–31, 133–6, 139
intervention 213–15, 216, 218–19, 223, 224, 225, 227 n. 20, 228
intuition and common-sense 1, 2, 3, 5, 7, 12, 13, 16, 17, 18, 20–1, 22, 23–4, 32, 48–9, 51, 54, 58–62, 64, 66, 68, 70, 76, 85, 87, 97–100, 102–3, 112, 118–19, 143–4, 153 n. 4, 159 n. 12, 182, 230–3, 234 (*see* conservatism, counterintuitiveness, parochialism)
Ismael, Jenann 11, 20, 22
isolation 121, 127

Jackendoff, Ray 104, 106
Jackson, Frank 63–4, 66
Jammer, Max 157 n. 8, 173 n. 22
Jeans, James
Physics and Philosophy 130
justification 8–9, 28–9, 73, 118, 123

Kant, Immanuel 9–10, 14, 35–6, 54 n. 9, 114, 118, 123, 139–40, 146, 158, 163, 182–8, 192–3, 196 (*see* neo-Kantians)
Kelvin, Lord (William Thomson) 189, 191, 192 n. 15
Kennedy, Jacqueline 104
Kepler, Johannes 35, 165
Kincaid, Harold 7, 21, 24
kind term 19–20, 200, 203–5
Kitcher, Philip 23, 90
Koyré, Alexandre 185
Kripke, Saul 159, 189, 190
Kuhn, Thomas 35, 114–15, 160–1, 164–5, 170, 185, 186, 188–9, 191–2
The Structure of Scientific Revolutions 114–15, 160–1, 185 n. 5

Ladyman, James 4, 7, 8, 9, 11, 13–14, 16, 17–18, 20, 22, 23, 29 n. 1, 39 n. 6, 41 n. 7, 47, 52 n. 4, 68 n. 31, 81–94, 97, 234
Every Thing Must Go: Metaphysics Naturalized 4, 9, 13, 17–18, 22, 29 n. 1, 41 n. 7, 47, 52 n. 4, 68 n. 31, 81–95, 97, 108–48, 143, 234
Lakatos, Imre 116
laws of nature 68–70, 74–5, 108–9, 117, 128, 140, 144, 146, 183–6, 193–4, 208–12, 227–8
and causal information 212–20
Leibniz, Gottfried 175
Lewis, David 4, 56–8, 102–3, 159
Lewis, Stephanie 102–3
Lewy, Hans 171
Linnaeus 100, 101
logical depth 108–9
logical empiricism 27, 51, 141, 161 n. 13, 184–5
logical positivism 113–14, 118–19
Lorentz, Henrick 162, 194–5
Lowe, Jonathan:
Survey of Metaphysics 4

Mäki, Uskali 121
Macarthur, David 2–3
Mach, Ernst 11, 152, 153 n. 3, 175, 183, 186, 192–3
Maddy, Penelope
Second Philosophy 4–5, 9, 13, 14–15, 179 n. 27, 205 n. 7
manifest image 17, 24, 96–106 (*see* folk theories, scientific image)
Manin, Yuri 156 n. 7
Manley, David 5
many worlds (*see* Everett, Hugh)
matter theory 194–6
Maudlin, Tim 29 n. 1, 47, 57 n. 11
Maxwell, James Clerk 162, 185, 186, 194–5
McDowell, John 3
Melnyk, Andrew 16–17, 20, 23
melodies 101, 103
meme 106
Mercator 100
metaphysics
analytic 1, 3–7, 12–13, 17, 23–4, 29–30, 32–3, 51–8, 66–70, 75–6, 80–2, 88, 98–9, 118–19, 122, 143–4, 151–2, 158–9, 164, 175, 231–2, 234
constraints from science 2, 21–2, 23, 33, 40–3, 46–9, 55 (*see* primacy of physics constraint (PPC))
naturalized, incoherence of 37–44

naturalized, sanctioning relations for 40–9
simpliciter 31–7, 42, 44, 46
underdetermination by science 6, 15, 21–2, 39–40, 47–8, 55, 72, 94–5
Meyerson, Emile 185
microphysics 120, 143, 191, 192, 193, 195
miles 101–2
Minkowski, Hermann 183–4
Mlodinow, Leonard
The Grand Design 72 n. 35, 110
modus ponens 60
Montague grammars 178 n. 25
Monty Hall problem 60
multiverse (*see* Everett, Hugh)
Munitz, Milton 220 n. 13

Nagel, Ernest 63 n. 25, 161 n. 13
Nagel, Thomas 80 n. 3
natural ontological attitude 55 n. 10, 62, 73, 85
naturalism 1–26, 39–40, 198–207
varieties of 2–3, 14–15, 39–40
and holism 5–8
and normative epistemology 8
criticisms of 5–9, 19–20, 37–44
scientific philosophy 1–5, 11, 18–20, 151–81, 186–8
neo-Kantians 10, 35–6, 52, 136, 164 n. 17, 184–8, 205–6
Neurath, Otto 118
Neurath's boat 161 n. 13, 164, 171
new theory of reference 189–90
Newton, Isaac 35, 130, 131, 182–4, 186, 191, 192–4, 220
Newtonian physics 10–11, 35, 114, 119, 123, 125, 129, 172, 182–6, 190–4, 208, 211, 214 n. 6, 220
nominalism 53
non-Euclidean geometry (*see* geometry, Euclidean and non-Euclidean)
nouns, defective 101–2
numbers 101, 105, 128, 180 n. 28, 199
numerals 105

objectivity 109, 180 n. 29, 182, 187–8
correspondence 180 n. 29, 187–8
intersubjective agreement 11, 19, 180 n. 29, 187–8
in science 11, 18–19, 22, 84, 109, 113–15, 118, 119 n. 5, 141–2, 182, 184–5, 187–8
observables and unobservables 27–8, 44–5, 65–6, 67, 68–9, 71, 116–18, 127–8, 130 (*see* experiential distance)
Occam's razor 101
ontology 31–6, 51–8, 64–78, 100–6, 108–9, 111, 124–5, 128, 129–30, 131–9, 146
scientific ontology 16, 21–2, 55, 62, 75–6
speculative ontology 16, 21–2, 54–8, 63 n. 25, 67–8, 69–70, 75–6
ordinary language philosophy 100, 101
Overbye, Denis 209

paradigms 11, 12, 185, 188–9, 192
parochialism 108–9, 112, 116–17, 143–4
Pauli, Wolfgang 134
PayPal 104
Pearl, Judea 20–1, 22, 213–14, 221, 225, 226–8, 233–4
Causality 213–14, 221, 225, 227, 233–4
Peirce, Charles Sanders 141–6
perceptual abilities 59–60, 65, 70, 71
perceptual intuitions 59–61
pessimistic meta-induction 57, 76
physicalism 19–20, 56, 91–3, 179–80
Plato 144
Platonism 42–3, 68, 105
Poincaré, Henri 11, 152, 175, 186, 195
Popper, Karl 37–8, 115–16, 117, 118, 122–3, 124, 140–2
positivism 18, 22, 113–15, 118–19, 122–3, 141, 152–4, 157–64, 165, 171–3, 175, 177
Potrc, Matjaž 121
prediction 45, 46 n. 8, 116, 125, 127–8, 129, 137, 138–9, 142, 156, 158, 166–8, 212, 213, 216, 220
primacy of physics constraint (PPC) 91–3
principle of equivalence 10, 184, 185, 195
principle of least epistemic risk 58, 70–2
principle of mediocrity 117
principle of naturalistic closure (PNC) 89–91, 94
principle theory (Einstein) 195

problem of causal antecedents (*see* free will)
psychophysics (*see* Fechner, Gustav)
Putnam, Hilary 14, 35, 61 n. 23, 189–90, 196

quantum physics 13, 17–18, 21–2, 43, 55, 57, 68–9, 80 n. 5, 103, 112, 113, 119, 120–1, 122, 123–6, 129, 131–8, 143, 146, 195, 208 n. 2
  Bell inequality 136
  Bohr's view of (*see* Bohr, Niels)
  Bohm interpretation of (*see* Bohm, David)
  Copenhagen interpretation 112, 130, 134, 136
  entanglement 57, 80 n. 5, 136–8, 143
  Everett interpretation of (*see* Everett, Hugh)
  formalism about 22, 126, 129–30, 133–5, 139
  Ghirardi–Rimini–Weber theory (GRW) 132, 134, 136
  locality 132, 136
  measurement problem 125–6, 132–7
  quantum field theory 69, 135, 137, 143, 195
  wave function 125, 130, 132–4, 135–6, 137–9
  collapse 125, 132–4, 135–6, 137 (*see* Copenhagen interpretation)
Quine, Willard Van Orman 5–8, 9, 11, 12, 15, 18, 39, 72, 73–4, 101–2, 105, 114, 119 n. 5, 153 n. 5, 155, 158, 159–60, 161, 162, 164, 175, 188–9, 198, 205–6
  holism 5–8, 11, 18, 114, 159–60, 188
  ontological relativity 17, 72, 105
  web of beliefs 6, 9, 12, 188–9
  Word and Object 101–2

real patterns 13–14, 108–9, 121–2, 128, 130, 137–8, 146, 148
realism and anti-realism 52–4
  causal 226–8
  metaphysical 35–6
  scientific 27–9, 39, 45, 71–4, 114, 119, 123–39, 146, 147, 180, 187–92, 196
  selective 39 n. 6, 73–4
  structural (*see* structural realism)
reductionism and anti-reductionism 17, 38, 51 n. 1, 63–4, 100, 108–9, 120–1, 143, 145–6
Reichenbach, Hans 2, 10, 118, 123, 152, 158, 184, 185 n. 4, 193–4
  coordinating principles 193–4
Reid, Thomas 153 n. 4, 175
relativity 10–11, 19, 47, 114, 118, 129, 136–8, 155, 160–1, 162–3, 183–6, 193–5, 220 n. 13
religion 52 n. 5, 52 n. 6, 67
representation 65–6
risk (*see* epistemic risk, inductive risk)
robot 97–8
Ross, Don 4, 7, 8, 9, 11, 13–14, 16, 17–18, 20, 22, 23, 29 n. 1, 41 n. 7, 47, 52 n. 4, 68 n. 31, 81–94, 97, 234
  Every Thing Must Go: Metaphysics Naturalized 4, 9, 13, 17–18, 22, 29 n. 1, 41 n. 7, 47, 52 n. 4, 68 n. 31, 81–95, 97, 108–48, 143, 234
Rosser, John Barkley 58 n. 15
Russell, Bertrand 11, 18, 152–3, 157, 177, 187, 211–12, 213
  Problems of Philosophy 11, 152

sakes 101, 102, 103
Saunders, Simon 126
scaffolding 220–1, 224, 228
scale invariance 16, 55, 68–70, 76, 108, 116–17, 120, 121, 135, 163
Scheines, Richard 20, 213
Schlick, Moritz 152, 153 n. 5, 184, 185 n. 4, 187
scientific image 99, 100, 103 (*see* manifest image)
scientific progress 37–9, 119, 129, 147
scientific revolutions 115, 160–1, 185–6, 188–9, 194–5
scope of models 223–7, 230
Searle, John 4
second philosophy (*see* Maddy, Penelope)
Sellars, Wilfrid 17, 100 (*see* manifest image)
semantics 155–64, 168–78
  semantic externalism 61–2, 190
Shaffer, Jonathan 80 n. 5
simultaneity 184–5
smithereens 102
Smolin, Lee 110 n. 2

Sober, Elliott 7
software 106
space and time 35, 131, 56, 69, 132, 137–8, 182–5, 192–3, 195–6, 204, 231–2, 233
  absolute 35, 131, 182–3, 184, 192–3
  spatiotemporal structure 185, 193–5
special sciences 13–14, 16, 90–3, 108, 122
Spirtes, Peter 20, 213
Spurrett, David 81–94
statistics 8, 17–18, 111, 134, 137, 139–48
stochasticity, irreducible 17–18, 125, 139–40, 144, 146
Strawson, Galen 233
structural realism 13, 133, 135–6, 147–8
Sturm, Charles-François 190, 200 n. 2, 201–5
susceptibility (*see* epistemic risk)

Tarski, Alfred 177
Taylor, Richard 209–10
theory change 39, 109, 114–15, 119, 123, 129, 185–6
theory facade 11–12, 190–4
thought experiments 70, 155, 159–61, 162, 164, 173–5
types and tokens 104–6

underdetermination (*see* metaphysics, underdetermination by science)
unification 8, 14, 16, 20–1, 23, 81–3, 89–91, 94, 108, 110, 113, 116, 121–2, 124, 125, 127–8, 137
  and disunity theses 121–2
utility functions 124, 128, 130, 138–9

van Fraassen, Bas 51 n. 2, 109–10, 130
  The Empirical Stance 51 n. 2, 109, 130
variables, exogenous and endogenous 221–2, 224–6, 227 n. 20, 226–30, 232
verificationism 86–9, 140–1
voices 103, 105–6
voluntarism 30–1
von Helmholtz (*see* Helmholtz, Hermann)
von Neumann, John 134

Wallace, David 126 n. 6, 135, 136
Watt, James 176
Weebles 104–5
Weinberg, Stephen
  Dreams of a Final Theory 110
Williams, Michael 8, 9, 15,
Wilson, Mark 5, 7, 8, 9, 11–13, 15, 18–20, 23–4, 119, 129, 147, 186–96
  Wandering Significance 9, 11–13, 15, 19, 119, 129, 147, 151, 152–3, 157 n. 10, 158–9, 164, 165, 166 n. 18, 170–1, 172, 174, 175, 176, 186–96
Wittgenstein, Ludwig 111, 157, 175–8
  Philosophical Investigations 157
  Tractatus Logico-Philosophicus 111
Woodhouse, Richard 170
Woodward, James
  Making Things Happen 213
words (as entities in an ontology) 105–6
Wright, Crispin 188 n. 9, 188 n. 10